The language of civil engineering research articles ~ a corpus-based approach

Alex Gilmore
Department of Civil Engineering
University of Tokyo, Japan

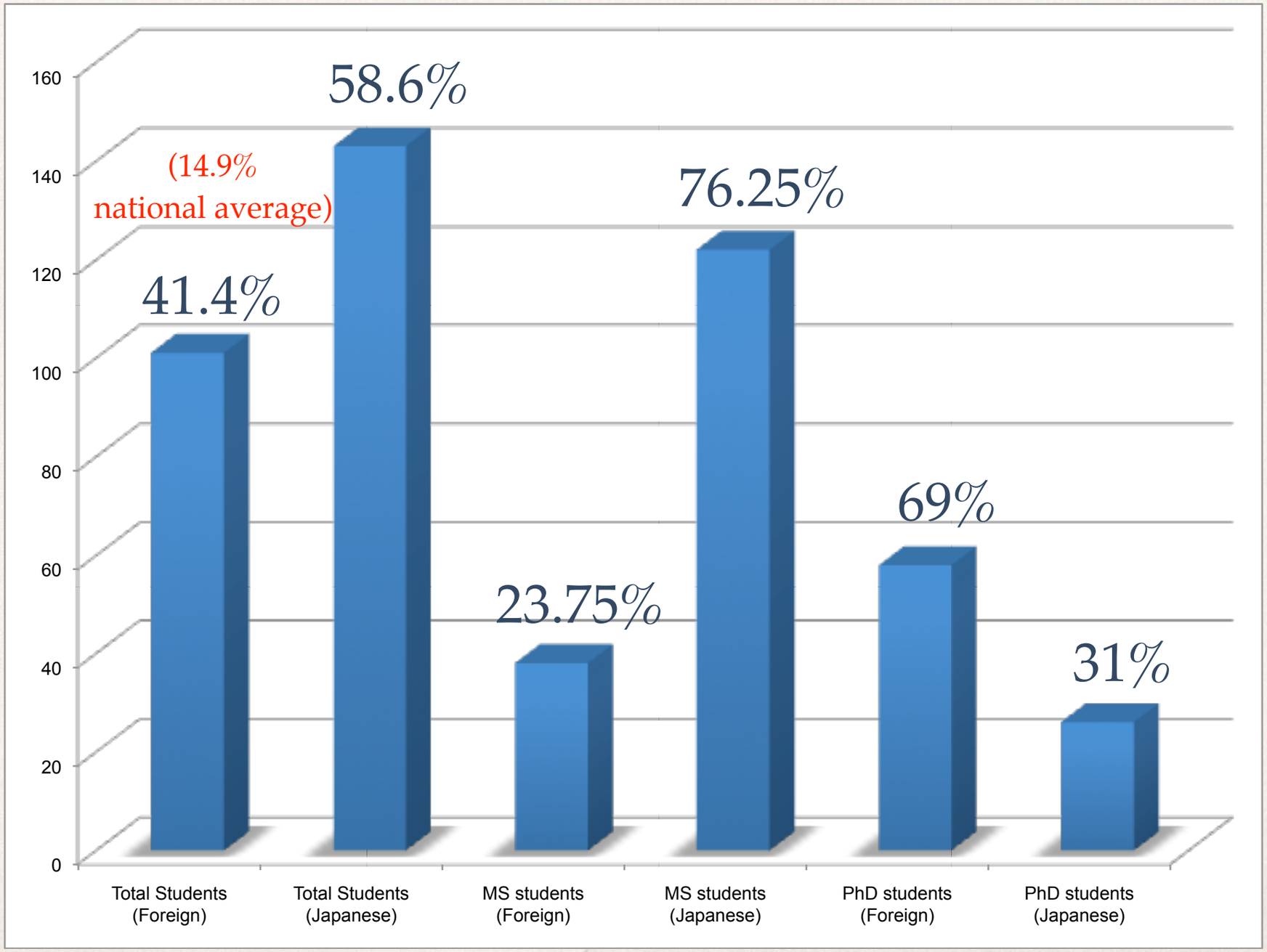
Aston University, September, 2015

Department of Civil Engineering, University of Tokyo

Department of Civil Engineering

1. Infrastructure Development & Construction Management
2. Regional Planning & Surveying
3. Transportation Engineering & Planning
4. River & Environmental Engineering
5. Coastal & Ocean Engineering
6. Hydrology & Water Resources Engineering
7. Geotechnical Engineering
8. Concrete & Construction Engineering
9. Earthquake & Disaster Mitigation Engineering
10. Mechanics & Structures
11. International Projects

Postgraduate student population - Department of Civil Engineering, University of Tokyo



Incoming students: Where are they from?

Why create a specialised corpus?

- Large variation between different academic disciplines in terms of word frequencies, collocational patterns & rhetorical moves: e.g. 4-word lexical bundles from fields of Biology, Electrical Engineering, Applied Linguistics & Business Studies >50% unique (Hyland 2008)

- Specialised corpora a good starting point for design of ESP materials - an area where publishers are unwilling to invest resources due to the restricted audience

SCCERA (Specialized Corpus of Civil Engineering Research Articles)

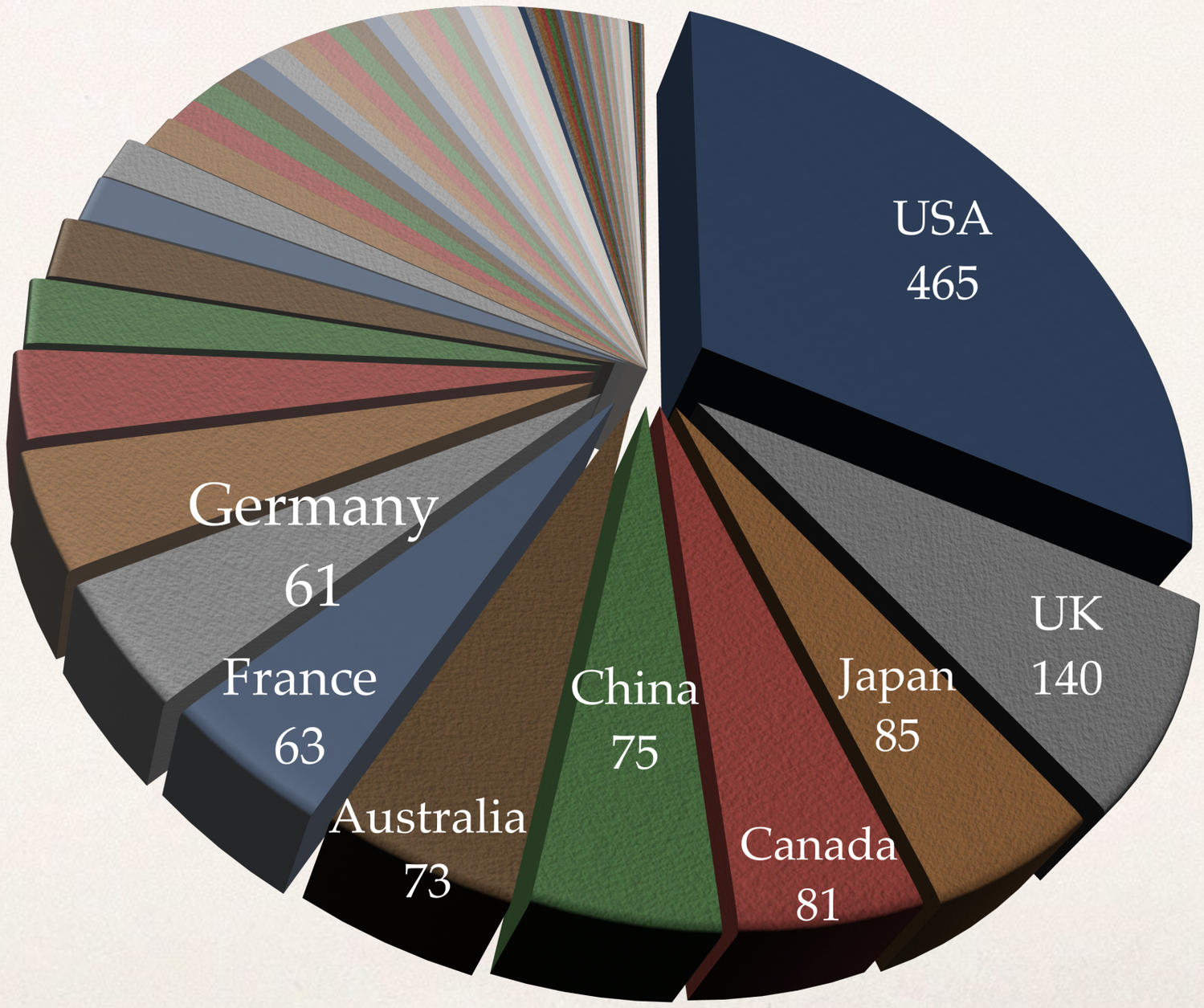
- **Phase 1**: Consultation with corpus linguists & civil engineers on the design & make-up of SCCERA (balanced & representative)
- **Phase 2**: Construction of SCCERA
- **Phase 3**: Quantitative & qualitative analysis of the corpus
- **Phase 4**: Exploring pedagogic applications of the corpus
- **Phase 5**: Dissemination of research results

SCCERA characteristics

- Total size: ~ 8 million words
- 45 journals (43 cited in SCI Expanded or SSCI)
- 1,100 research articles (most cited or downloaded)
- Year of publication: Range = 1989 - 2014; Mean = 2009
- 3,807 contributing authors (average of 3.46 authors per article)
- 1,598 participating institutions from 80 countries

Participating institutions by country (N = 80)

Participating institutions by country (N = 80)



Research questions

- What are the most frequently occurring keywords in civil engineering RAs?
- How are keywords from SCCERA distributed across established external wordlists - General Service List (West, 1953) & Academic Word List (Coxhead, 2000)?
- How are keywords from SCCERA distributed across the sub-divisions of civil engineering?
- What are the most frequent 3 - 6-word bundles in civil engineering RAs?
- How can this data inform materials design in an ESP context?

Keywords analysis (BNC reference corpus)

- Highlights words which are unusually frequent/infrequent compared to a reference corpus & helps to characterize the 'aboutness' of a genre

- Potentially more useful for materials design than raw word frequency data

- SCCERA keywords focus on materials; observing, measuring & reporting on their physical properties or behaviours

- Large degree of nominalisation (analysis, behaviour, distribution, etc), shifting the focus from an agent to an object or concept

Top 50 keywords in SCCERA

1	et al	26	temperature
2	fig	27	measured
3	model	28	behavior
4	data	29	coefficient
5	equation	30	ratio
6	results	31	spatial
7	values	32	variables
8	models	33	distribution
9	flow	34	strain
10	concrete	35	method
11	table	36	parameter
12	shear	37	measurements
13	using	38	shown
14	wave	39	earthquake
15	figure	40	value
16	surface	41	density
17	parameters	42	average
18	water	43	respectively
19	analysis	44	precipitation
20	eq	45	displacement
21	soil	46	effects
22	based	47	cement
23	stress	48	climate
24	observed	49	maximum
25	velocity	50	project

Top 40 keywords in SCCERA

et. al	table	soil	ratio
fig	shear	based	spatial
model	using	stress	variables
data	wave	the	distribution
equation	figure	observed	strain
results	surface	velocity	method
values	parameters	temperature	parameter
models	water	measured	measurements
flow	analysis	behavior	shown
concrete	eq	coefficient	earthquake

RANGE (Heatley &Nation, 1994) - Vocabulary profiling tool

	Families	Types	Tokens	Percent
K1 Words (1-1000):	99	134	**142**	**29.40%**
Function:	...	...	(13)	(2.69%)
Content:	...	...	(129)	(26.71%)
> Anglo-Sax =Not Greco-Lat/Fr Cog:	...	...	(37)	(7.66%)
K2 Words (1001-2000):	41	53	**53**	**10.97%**
> Anglo-Sax:	...	...	(20)	(4.14%)
1k+2k			...	(40.37%)
AWL Words (academic):	97	145	**145**	**30.02%**
> Anglo-Sax:	...	...	(4)	(0.83%)
Off-List Words:	?	143	**143**	**29.61%**
	237+?	475	483	100%

Current profile	
%	Cumul.
29.40	29.40
10.97	40.37
30.02	70.39
29.61	100.00

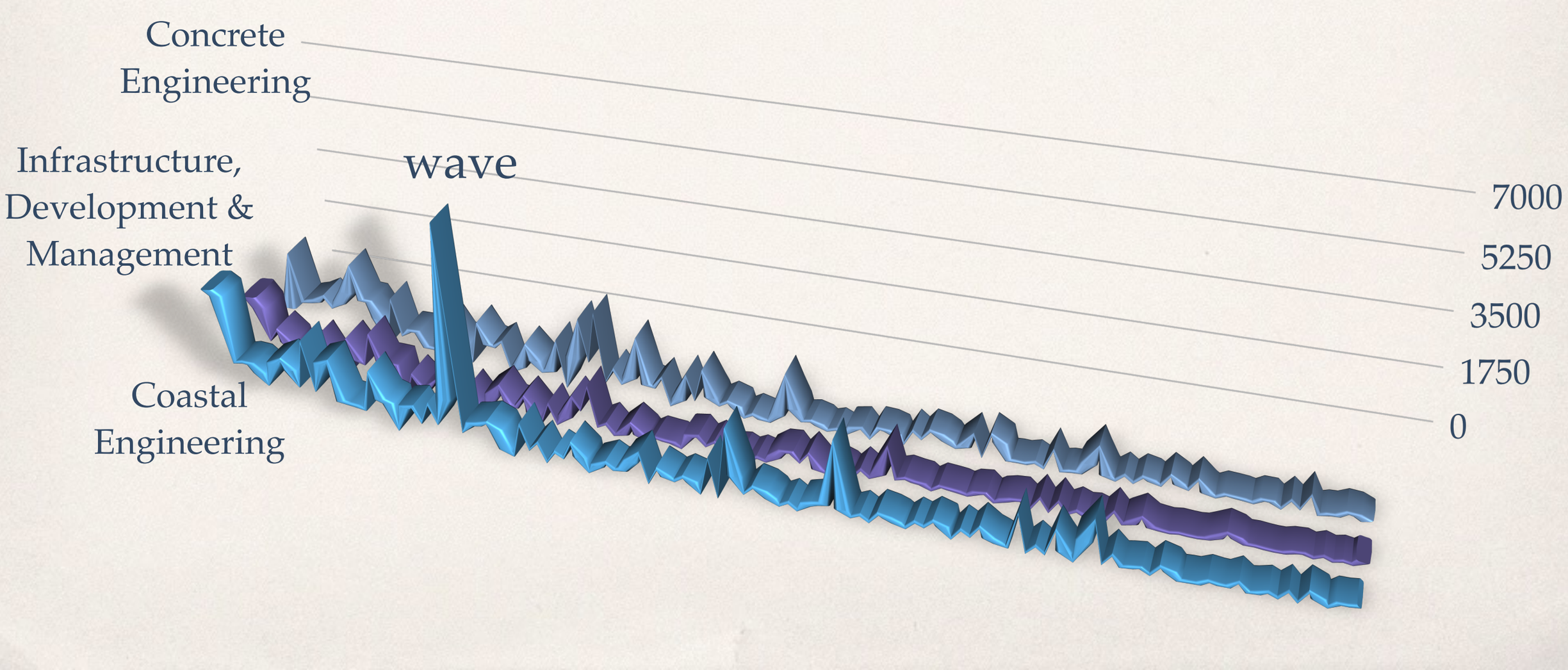
http://www.lextutor.ca (Tom Cobb)

Distribution of keyword families across GSL, AWL & ‘off-list’

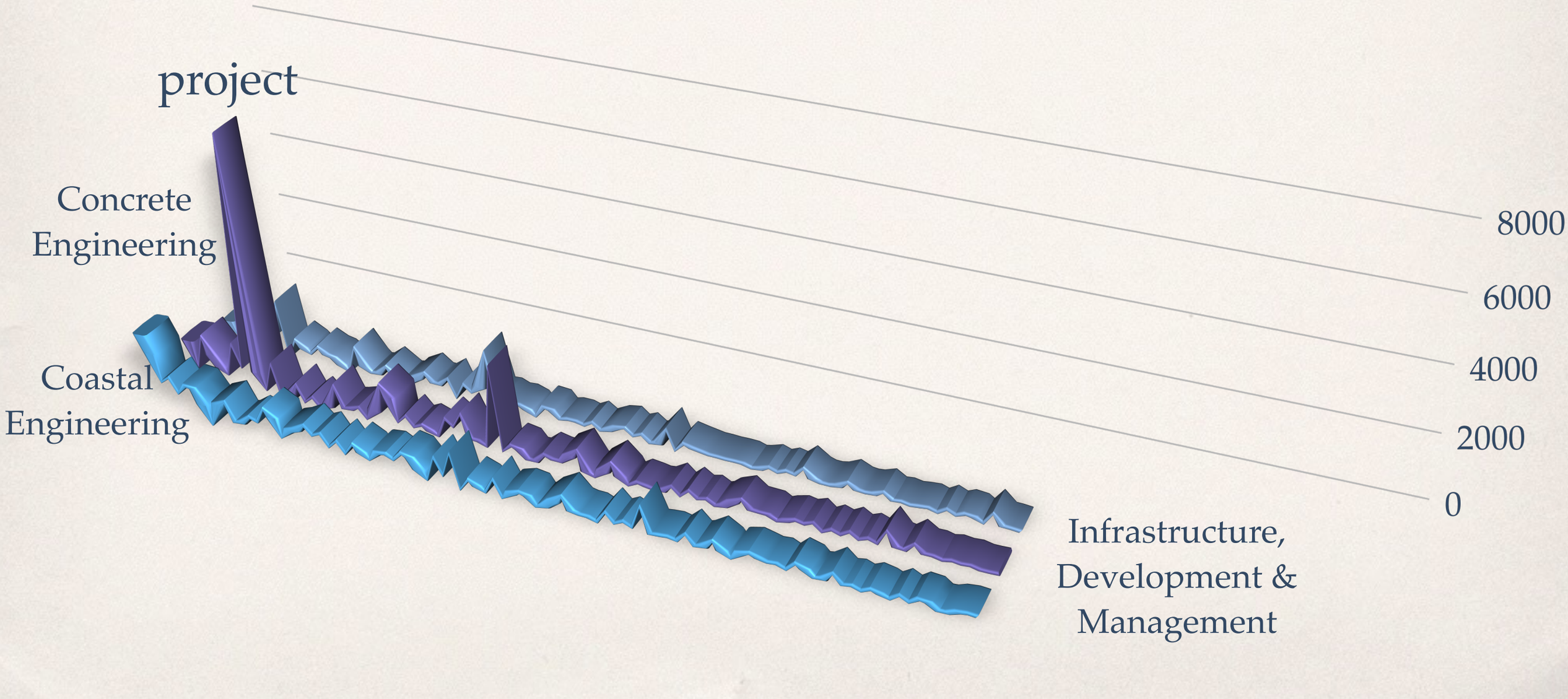
Distribution of keyword families across GSL, AWL & 'off-list'

"Taken together, the first 2,000 words in West's (1953) GSL and the word families in the AWL account for approximately 86% of the Academic Corpus" (Coxhead 2000: 222)

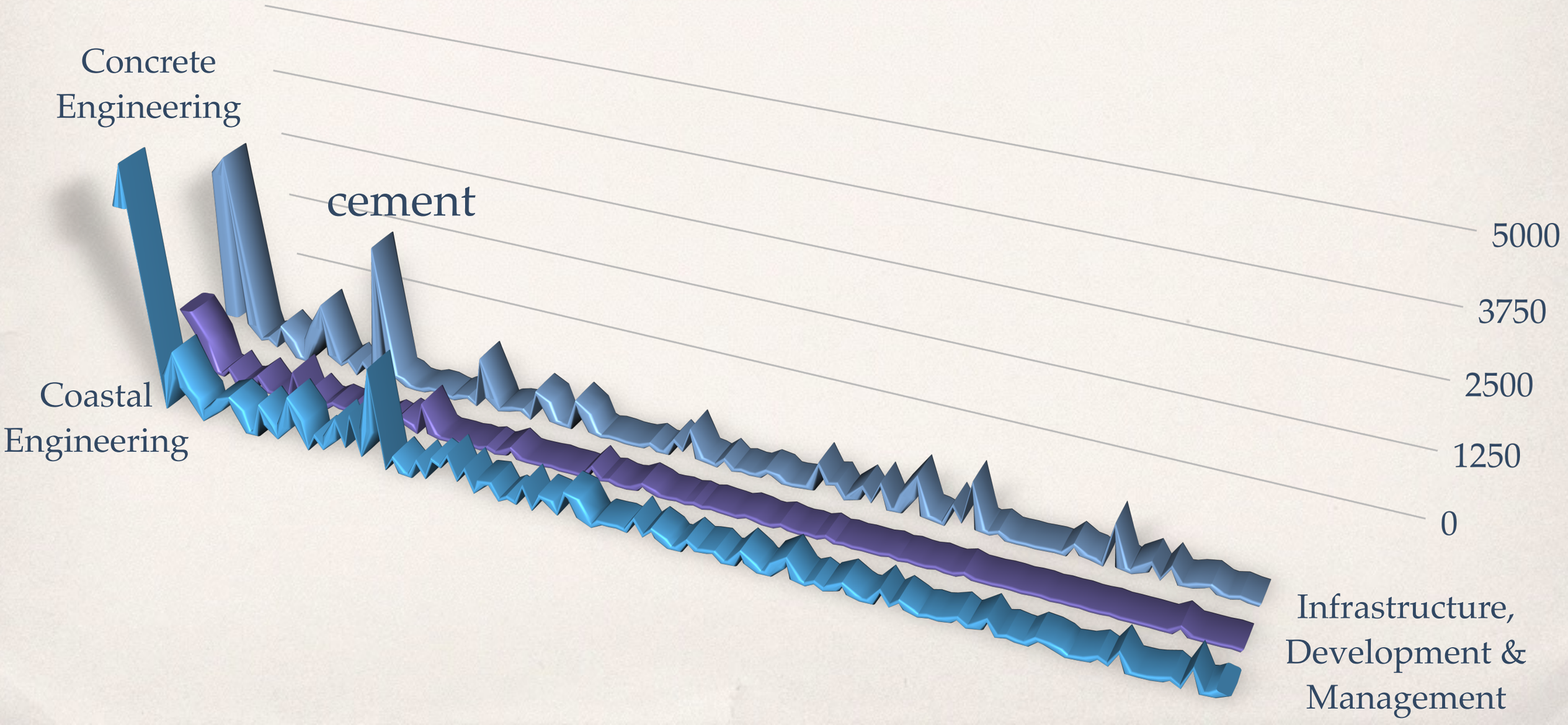
Distribution of keyword families across SCCERA: GSL



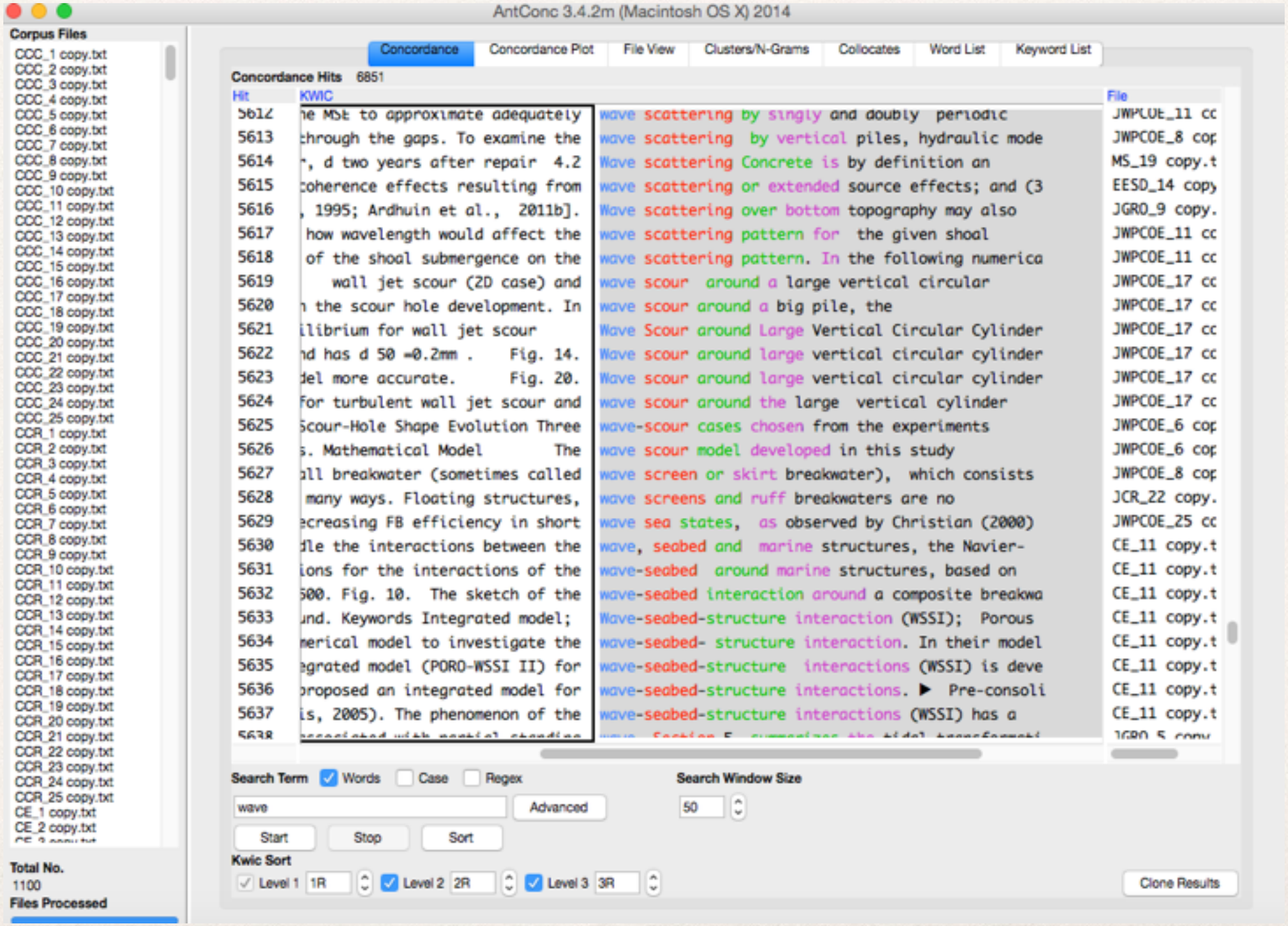
Distribution of keyword families across SCCERA: AWL



Distribution of keyword families across SCCERA: Offlist



AntConc - *Wave* collocates



AntConc - *Wave* collocates (80+)

Collocate	Hits
height	813
velocity	279
period	263
break	249
energy	174
model	148
condition	147
propogation	145
runup	115

Lexical bundles

- Defined as “the most frequently occurring lexical sequences in a register” (Biber, Conrad & Cortes 2004)

- “lexical bundles are crucially important for the construction of discourse in all university registers” (Biber 2006: 174)

Lexical bundles are…

- Extremely common
- Usually not idiomatic in meaning
- Usually not perceptually salient
- Usually not complete structural units, but instead tend to bridge 2 phrases (e.g. ~5% for academic writing)
- Function as ‘discourse frames’ for the expression of new information - they don’t express new propositional meaning themselves

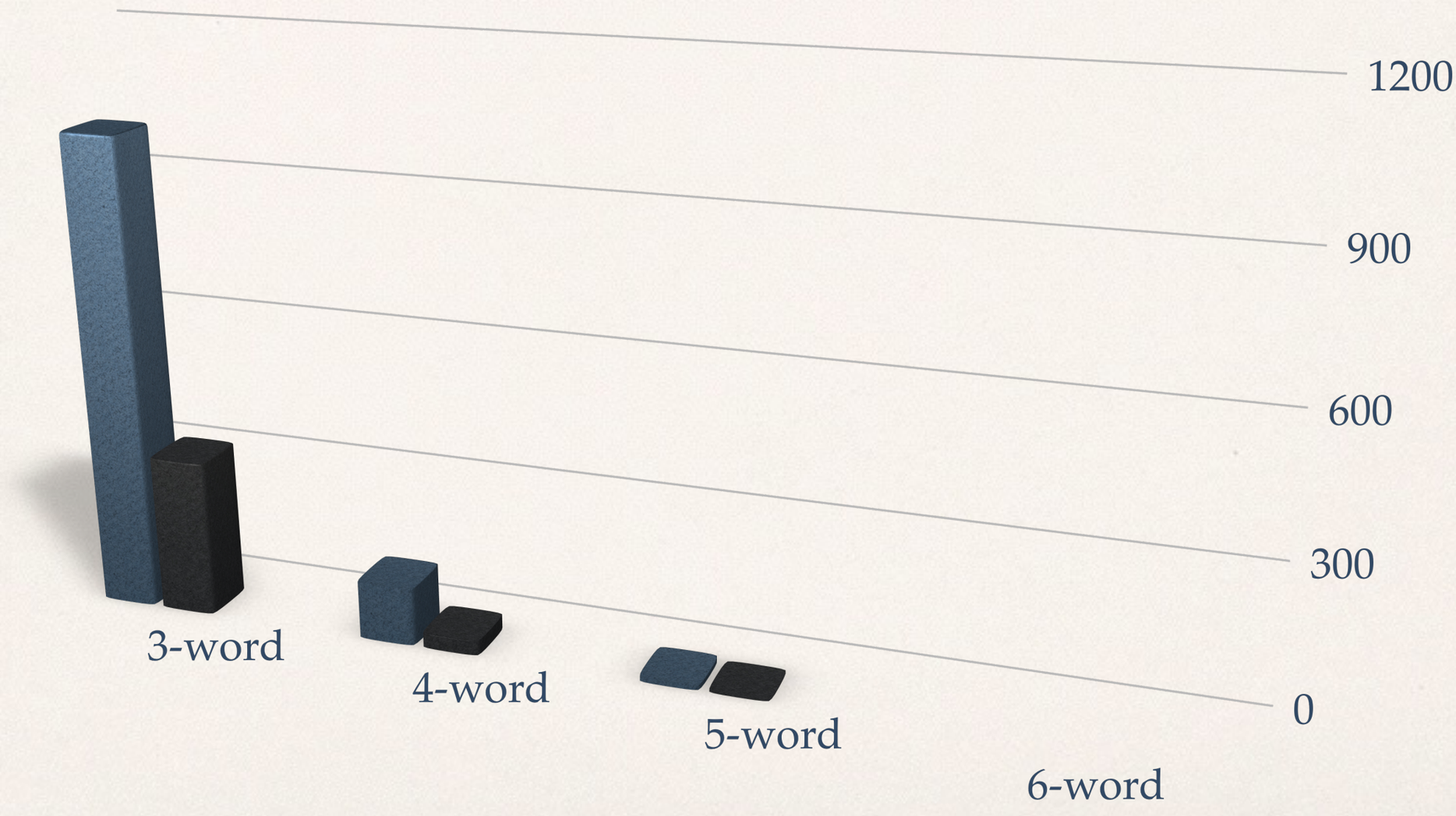
(Biber & Barbieri 2007: 269/70)

Lexical bundles (formulaic language

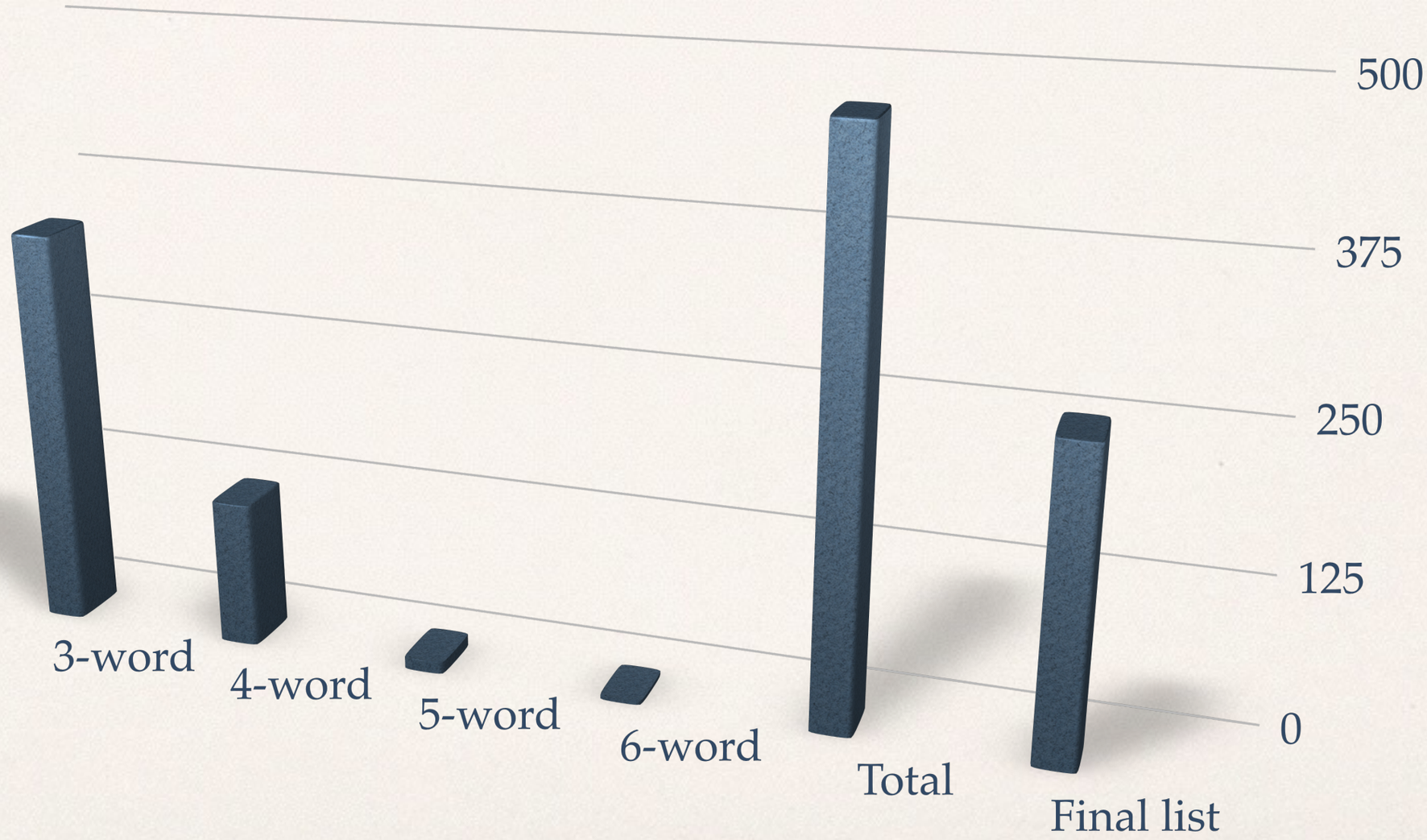
"Given the importance of formulaic language, it can be argued that it needs to be part of language syllabuses […] Unfortunately, this is not the case. A perusal of almost any EFL/ESL textbook or test yields a paucity of formulaic sequences targeted for explicit attention/noticing, and even for those that do occur, there does not seem to be much principled basis for selection"

(Martinez & Schmitt 2012: 301)

Lexical bundles in SCCERA (20 or 40 occurrences per million cut-off)



Lexical bundles in SCCERA



Selecting lexical bundles for pedagogical applications

- Start with longest bundles (6-word), e.g. *it should be noted that the*

- Work through list from longest to shortest (6-word ➡ 3-word), using key words (e.g. *noted*) to search for related bundles: *be noted that; should be noted*

- Use KWIC function in a concordancer (e.g. AntConc) to decide if bundles are fragments of longer bundles or not

Selecting lexical bundles - examples in the decision-making process

- *should be noted* - excluded from list (KWIC search of SCCERA using AntConc indicates it is a fragment of *it should be noted that*)

  should be noted: total hits = 374

  Right sorted: *should be noted that* = 339 hits (90.6%)

  Left sorted: *it should be noted* = 364 (97.3%)

- *in the form* - included in final list (not solely a fragment of *in the form of,* but also *in the form + eqn*)

  in the form: total hits = 395

  Right sorted: *in the form of* = 335 hits (84.8%)

Selecting lexical bundles - examples in the decision-making process

- *should l* ntConc indicat

  should

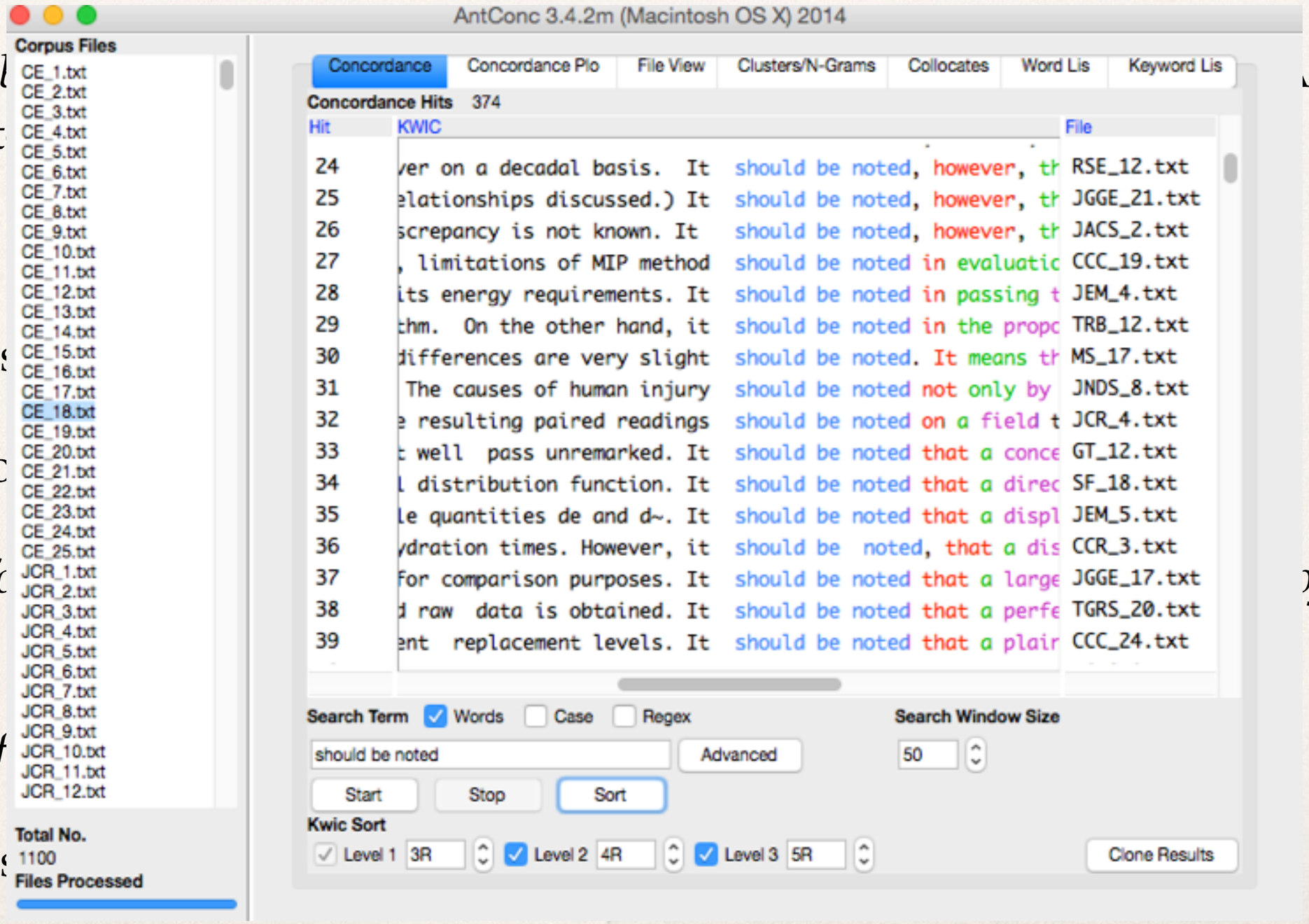
  Right s

  Left sc

- *in the fo* of, but also *in*

  in the f

  Right s



Selecting lexical bundles - examples in the decision-making process

- *should be noted* - excluded from list (KWIC search of SCCERA using AntConc indicates it is a fragment of *it should be noted that*)

  should be noted: total hits = 374

  Right sorted: *should be noted that* = 339 hits (90.6%)

  Left sorted: *it should be noted* = 364 (97.3%)

- *in the form* - included in final list (not solely a fragment of *in the form of,* but also *in the form + eqn*)

  in the form: total hits = 395

  Right sorted: *in the form of* = 335 hits (84.8%)

Selecting lexical bundles - examples in the decision-making process

- *should b* ... AntConc indicate

  should ...

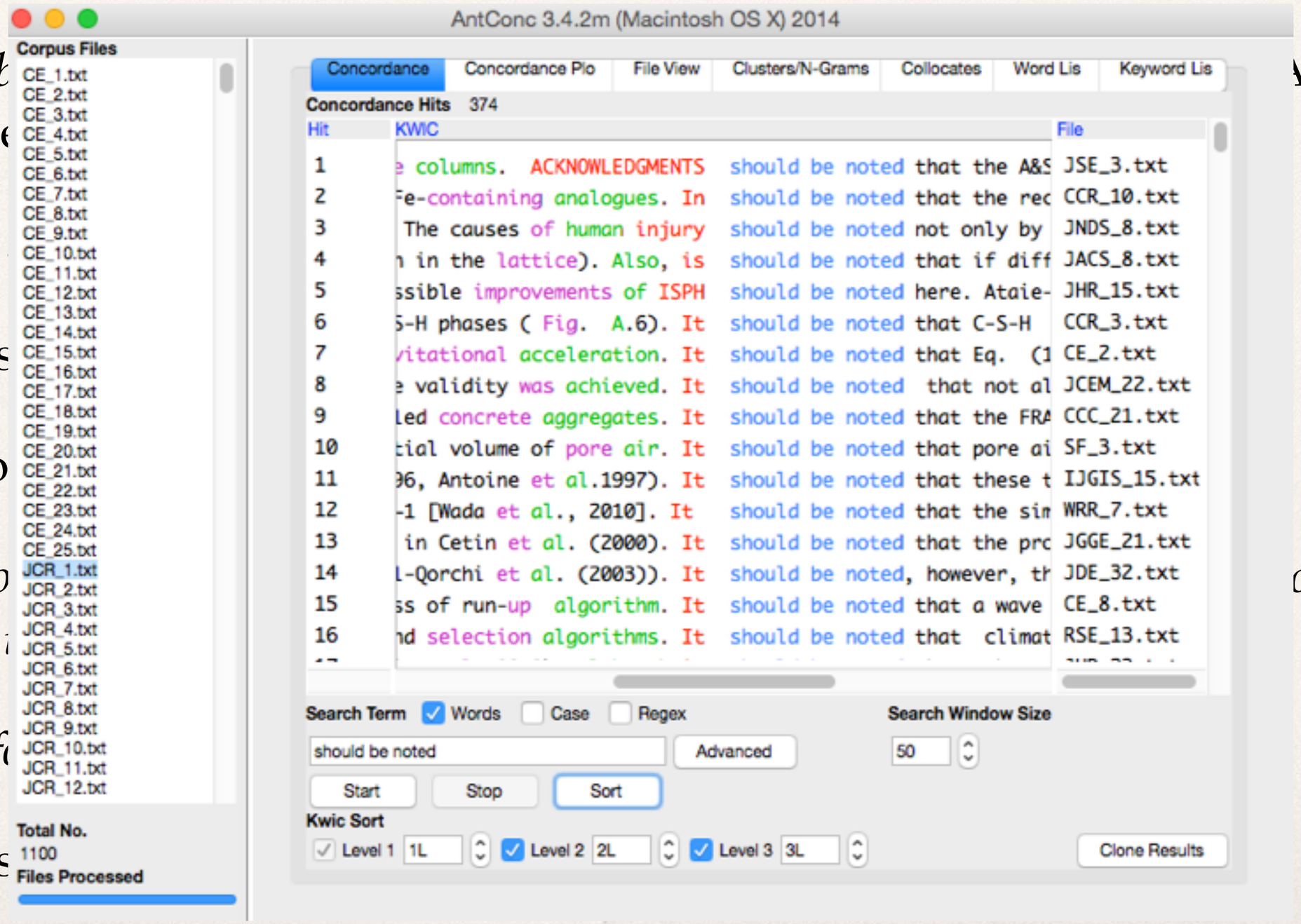
  Right s...

  Left so...

- *in the fo* ... *of*, but also *in* ...

  in the f...

  Right s...



Selecting lexical bundles - examples in the decision-making process

- *should be noted* - excluded from list (KWIC search of SCCERA using AntConc indicates it is a fragment of *it should be noted that*)

  should be noted: total hits = 374

  Right sorted: *should be noted that* = 339 hits (90.6%)

  Left sorted: *it should be noted* = 364 (97.3%)

- *in the form* - included in final list (not solely a fragment of *in the form of,* but also *in the form + eqn*)

  in the form: total hits = 395

  Right sorted: *in the form of* = 335 hits (84.8%)

Selecting lexical bundles - examples in the decision-making process

- *should be* AntConc indicate

  should b

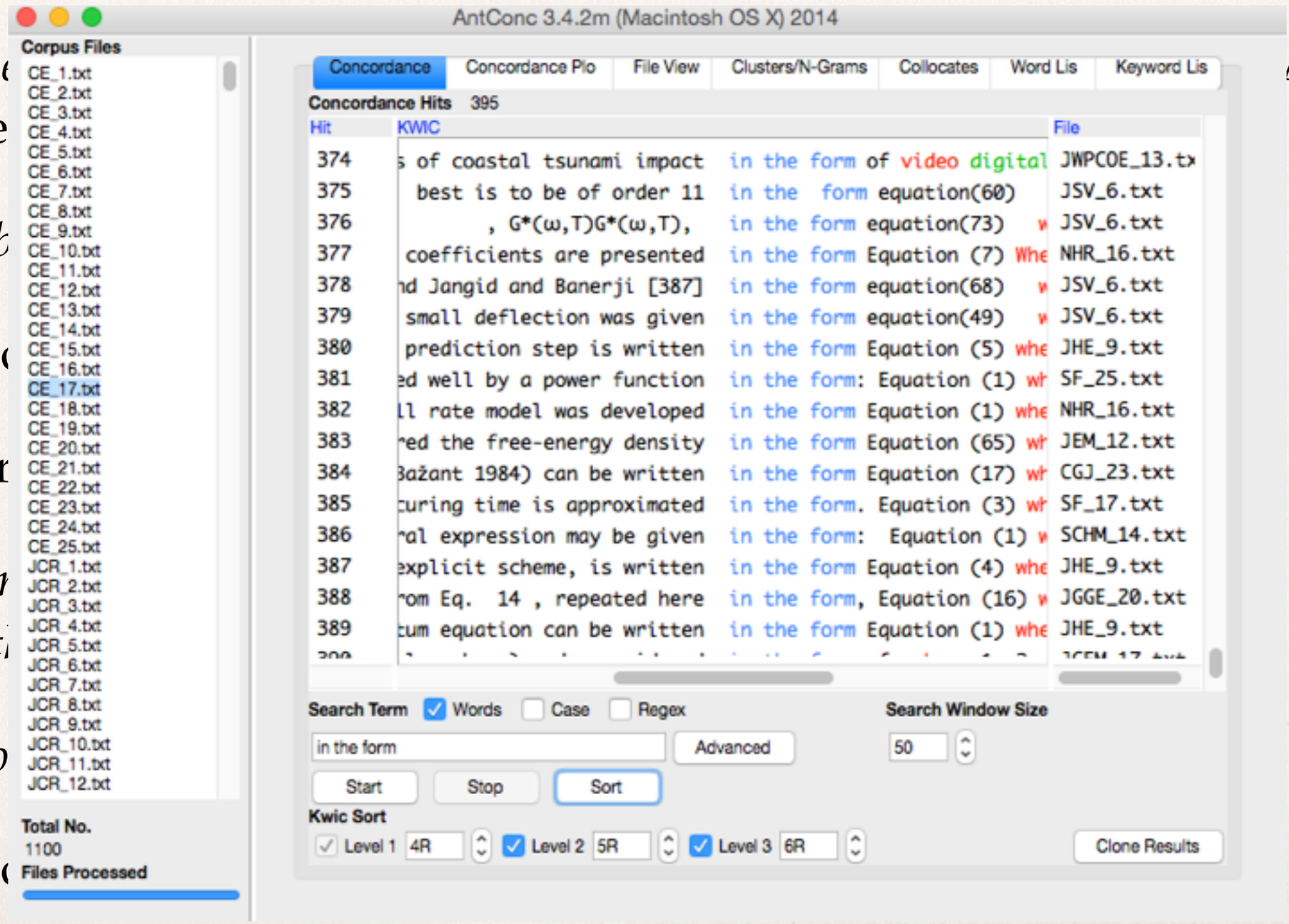
  Right s

  Left sor

- *in the for* *of*, but also *in t*

  in the fo

  Right s



Selecting lexical bundles - examples in the decision-making process

- *should be noted* - excluded from list (KWIC search of SCCERA using AntConc indicates it is a fragment of *it should be noted that*)

  should be noted: total hits = 374

  Right sorted: *should be noted that* = 339 hits (90.6%)

  Left sorted: *it should be noted* = 364 (97.3%)

- *in the form* - included in final list (not solely a fragment of *in the form of,* but also *in the form + eqn*)

  in the form: total hits = 395

  Right sorted: *in the form of* = 335 hits (84.8%)

Selecting lexical bundles - examples in the decision-making process

- *should be noted* - excluded from list (KWIC search of SCCERA using AntConc

indi

sho

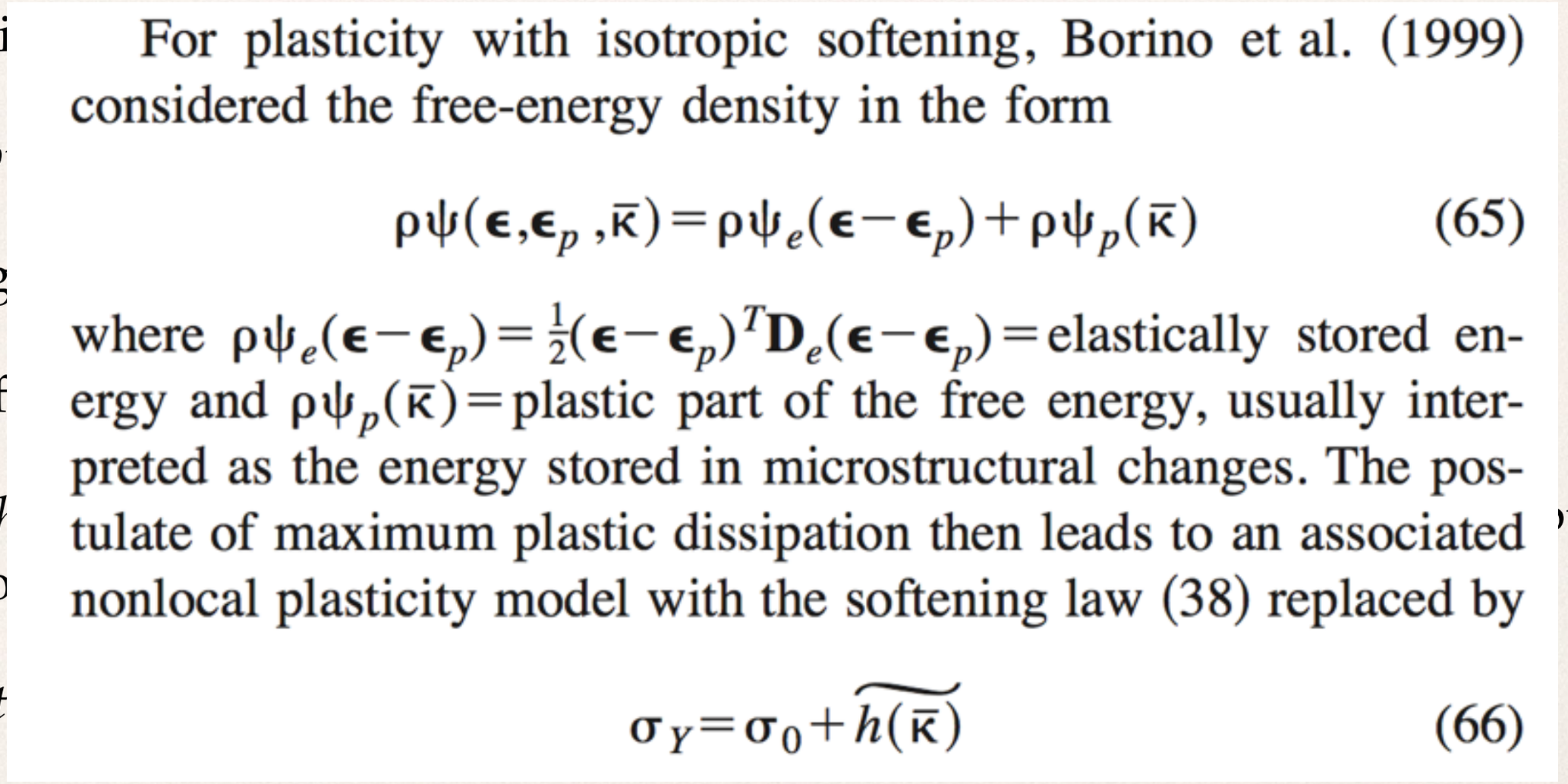
Rig

Lef

- *in th*

also

in t



For plasticity with isotropic softening, Borino et al. (1999) considered the free-energy density in the form

$$\rho\psi(\boldsymbol{\epsilon},\boldsymbol{\epsilon}_p,\bar{\kappa}) = \rho\psi_e(\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_p) + \rho\psi_p(\bar{\kappa}) \tag{65}$$

where $\rho\psi_e(\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_p) = \frac{1}{2}(\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_p)^T\mathbf{D}_e(\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_p)$ = elastically stored energy and $\rho\psi_p(\bar{\kappa})$ = plastic part of the free energy, usually interpreted as the energy stored in microstructural changes. The postulate of maximum plastic dissipation then leads to an associated nonlocal plasticity model with the softening law (38) replaced by

$$\sigma_Y = \sigma_0 + \widetilde{h(\bar{\kappa})} \tag{66}$$

ut

Right sorted: *in the form of* = 335 hits (84.8%)

Selecting lexical bundles - examples in the decision-making process

- *should be noted* - excluded from list (KWIC search of SCCERA using AntConc indicates it is a fragment of *it should be noted that*)

  should be noted: total hits = 374

  Right sorted: *should be noted that* = 339 hits (90.6%)

  Left sorted: *it should be noted* = 364 (97.3%)

- *in the form* - included in final list (not solely a fragment of *in the form of,* but also *in the form + eqn*)

  in the form: total hits = 395

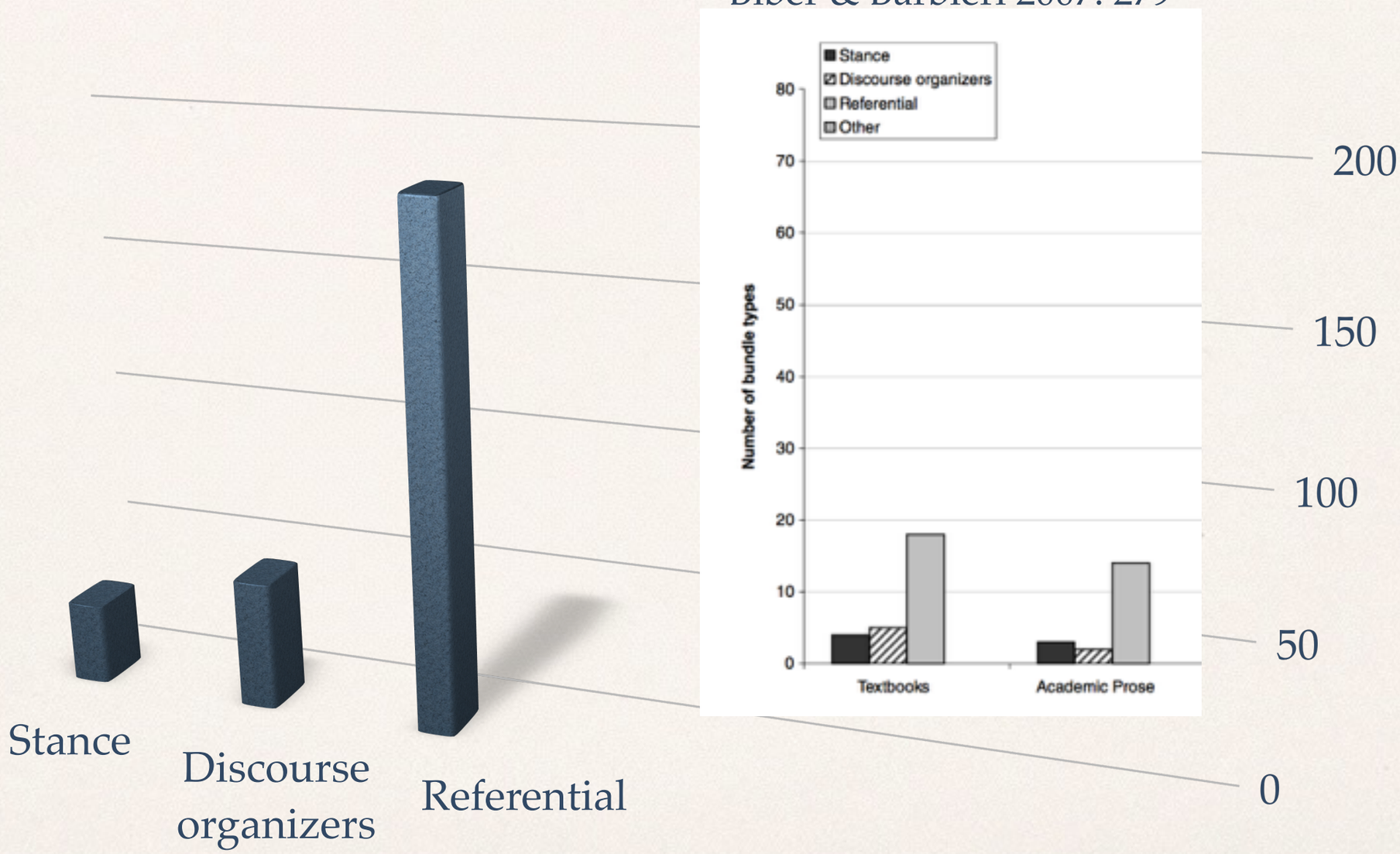
  Right sorted: *in the form of* = 335 hits (84.8%)

Typical patterns in clusters

- "Almost 30 per cent of the top 100 chains (types) have this pattern *PREP the N of the*. If we count closely related chains, with an indefinite article (*in the context of a*) or with embedded adjectives, such as *the ADJ N of the* (*the far side of the*), then, over 40 per cent of the top 100 chains have this pattern." (Stubbs & Barth 2003: 82)

- (34.7% in SCCERA)

Lexical bundles in SCCERA by type



Biber & Barbieri 2007: 279

Referential expressions

- Specifying attributes: *a little bit of; the size of the; the nature of the*
- Time/place/text deixis: *in the USA; at the end of the; as shown in Fig*

Conclusion

- Keywords from SCCERA (compared with GSL & AWL) provide the foundation for a more systematic approach to academic vocabulary development for civil engineering students

- Lexical bundles complement keyword lists in materials design by providing a 'phraseological profile' of a genre (Romer 2010: 27)

- Because lexical bundles are frequent but not perceptually salient, 'they might be good candidates for overt instruction' (Biber & Barbieri 2007: 284)

- "it is necessary to go beyond frequencies and complement the corpus-derived information with research tapping into their pedagogic relevance, with suggestions from teachers and students" (Flowerdew 2012: 268)

References

Biber, D. (2006). *University Language. A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins

Biber, D. & Barbieri, F. (2007). 'Lexical bundles in university spoken and written registers'. *English for Specific Purposes* 26: 263-286

Biber, D., Conrad, S. & Cortes, V. (2004). '*If you look at…*: Lexical bundles in university teaching and textbooks'. *Applied Linguistics* 25.3: 371-405

Coxhead, A. (2000). 'A new academic word list'. *TESOL Quarterly* 34.2: 213-238.

Flowerdew, L. (2012). *Corpora and language education*. Basingstoke: Palsgrave Macmillan

Heatley, A. & P. Nation (1994). 'Range', Victoria University of Wellington, NZ.

Hyland, K. (2008). 'As can be seen: Lexical bundles and disciplinary variation'. *English for Specific Purposes* 27: 4-21

Hyland, K. & Milton, J. (1997). 'Qualification and certainty in L1 and L2 students' writing'. *Journal of Second Language Writing* 6(2): 183-205

References

Martinez, R. & N. Schmitt (2012). 'A phrasal expressions list'. *Applied Linguistics* 33.3: 299-320

Pearson, J. (1998). *Terms in context.* Amsterdam: John Benjamins

Rea Rizzo, C. (2010). 'Getting on with corpus compilation: From theory to practice'. *ESP World* 1(9): 1-22

Römer, U. (2010). 'Using general and specialised corpora in English language teaching: Past, present and future'. In Campoy-Cubillo, Belles-Fortuno & Gea-Valor (Eds.) *Corpus-based approaches to English language teaching*. London: Continuum: 18-35

Scott, M. (2011). WordSmith Tools Version 6, Liverpool: Lexical Analysis Software.

Stubbs, M. & I. Barth (2003). 'Using recurrent phrases as text-type discriminators'. *Functions of Language* 10.1: 61-104

West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.

Lexical bundles

- Lexical bundles occurring at 20-40 times per million words or more considered (conservative - less conservative approaches: see Biber & Barbieri 2007)

- "an ESP perspective, considering each register on its own terms, is required to adequately describe the use of lexical bundles in the university context" (Biber & Barbieri 2007: 265)

- Three main functions of lexical bundles (Biber & Barbieri 2007: 270):

1. Expressions of stance (epistemic, desire, obligation, intention/prediction, ability)

2. Discourse organizers (topic introduction, topic elaboration/clarification, referential identification/focus)

3. Referential expressions (imprecision indicators, specifying attributes, time/place/text-deixis) - dominant function in written university registers

Conclusion

- Keywords from SCCERA (compared with GSL & AWL) provide the foundation for a systematic approach to academic vocabulary development for civil engineering students

- Word bundles???

- “A student in the university context will frequently encounter [word bundles]… failure to understand their textual and interpersonal functions will obviously influence her/his success in dealing with both spoken and written language situations” (Biber & Barbieri 2007: 284)

- “more work probably has to be put into the creation of reliable corpus-based language descriptions for learners and teachers, especially descriptions of specialized discourses” (Römer 2010: 29)

Conclusion

- “[…] some researchers have argued that factors such as perceptual salience and developmental readiness are more important than frequency for acquisition… If these findings are generalizable, they might indicate that the discourse functions of lexical bundles are not easily acquired: they are not perceptually salient, despite their high frequency. If so, lexical bundles might be good candidates for overt instruction.” (Biber & Barbieri 2007: 284)

- “it would seem sensible to combine our formulaic sequences into these wordlists in a way that would create a much more inclusive overall description of the most frequent (and therefore useful) lexical items of English, both individual and multi-word” (Martinez & Schmitt 2012: 304)

Lexical bundles can be deceptively transparent

- “It is not unreasonable to guess that L2 learners processing those expressions [in the PHRASE List] might therefore actually believe they understand them (if they identify them) simply because the individual words are so well known, making them, in Laufer’s (1989) terms, ‘deceptively transparent’ (Martinez & Schmitt 2012: 313)

Lexical bundles included in assessments of complexity?

- “There is also the potential […] for the development of an automated lexical profiling tool such as *Range* (Heatley and Nation 1994) that instead of only analyzing a text for individual words […] also carries out a ‘sweep’ for phrases […] to more accurately reflect its lexical complexity’ (Martinez & Schmitt 2012: 315/6)

Lexical bundles in SCCERA compared to PHRASE List

- Only **13.2%** of lexical bundles in SCCERA occur in Martinez & Schmitt's (2012) PHRASal Expressions List (PHRASE List) - 505 most frequent, non-transparent multi-word expressions in English (derived from BNC)

- Specialized corpus (SCCERA) vs general corpus (BNC)

- Transparency (receptive vs productive goals)

- 3- to 6-word bundles vs 2- to 4-word bundles

3-word clusters

based on the	with respect to	in the case
as well as	in this paper	there is a
the number of	one of the	the value of
in order to	in this study	the presence of
shown in fig	a function of	can be used
in terms of	the case of	the fact that
due to the	part of the	according to the
the effect of	a number of	as a result
the use of	the effects of	be used to
as shown in	the results of	the other hand

4-word clusters

in the case of	the results of the	it is important to
on the other hand	is shown in fig	it should be noted
as a function of	the size of the	in the context of
as shown in fig	are shown in fig	is assumed to be
as well as the	is based on the	the fact that the
can be used to	the end of the	should be noted that
on the basis of	at the end of	in the form of
with respect to the	the effect of the	it is possible to
in terms of the	at the same time	it can be seen
as a result of	in the united states	in this paper we

SCCERA journal list

Journal	Article Title	Article Code	Number Authors	Institution Countries	Year of Publication	Number Words
Coastal Engineering	Modelling storm impacts on beaches, dunes	CE_1	6	Netherlands; USA	2009	11,398
Coastal Engineering	Corrected Incompressible SPH method for a	CE_2	3	Japan; UK	2008	8,184
Coastal Engineering	44-year wave hindcast for the North East At	CE_3	3	Portugal; Spain	2008	2,817
Coastal Engineering	Increasing wave heights and extreme value	CE_4	3	USA	2010	10,302
Coastal Engineering	Modified Moving Particle Semi-implicit meth	CE_5	1	Japan	2009	11,452
Coastal Engineering	Numerical analysis of wave overtopping of r	CE_6	4	Spain	2008	7,208
Coastal Engineering	A 44-year high-resolution ocean and atmosp	CE_7	5	Spain; France	2008	8,829
Coastal Engineering	Simulation of nonlinear wave run-up with a	CE_8	2	Denmark	2008	7,499
Coastal Engineering	Beach Wizard: Nearshore bathymetry estim	CE_9	6	Netherlands; USA; Chile	2008	8,286
Coastal Engineering	Efficient computation of surf zone waves usi	CE_10	2	Netherlands	2008	6,532
Coastal Engineering	An integrated model for the wave-induced s	CE_11	4	UK; China; USA	2013	8,876
Coastal Engineering	Statistical simulation of wave climate and es	CE_12	4	Australia	2008	8,888
Coastal Engineering	Hindcast of the wave conditions along the w	CE_13	3	Portugal	2008	5,480
Coastal Engineering	Laboratory and numerical studies of wave d	CE_14	3	USA	2009	6,662
Coastal Engineering	A probabilistic methodology to estimate futu	CE_15	3	UK	2008	6,178
Coastal Engineering	Run-up of tsunamis and long waves in term	CE_16	2	Denmark	2008	8,813
Coastal Engineering	Measurement of wave-by-wave bed-levels i	CE_17	3	Australia; UK	2008	1,648
Coastal Engineering	The morphological response of a nearshore	CE_18	4	Netherlands; USA	2008	7,938
Coastal Engineering	Direct bed shear stress measurements in bc	CE_19	4	Australia; UK	2009	9,281
Coastal Engineering	Two-dimensional time dependent hurricane	CE_20	7	Netherlands; USA	2010	9,300
Coastal Engineering	Modeling hurricane waves and storm surge	CE_21	10	Netherlands; USA	2011	9,528
Coastal Engineering	On the evolution and run-up of breaking sol	CE_22	4	Taiwan	2008	7,007
Coastal Engineering	Morphodynamic responses to the deep wate	CE_23	3	China	2009	7,962
Coastal Engineering	Large-scale dune erosion tests to study the	CE_24	5	Netherlands	2008	5,852
Coastal Engineering	Wave boundary layer over a stone-covered	CE_25	4	Denmark	2008	11,155
J. of Coastal Research	The Role of Remote Sensing in Predicting ar	JCR_1	1	USA	2009	7,568
J. of Coastal Research	Shoreline Definition and Detection: A Revie	JCR_2	2	Australia	2005	5,895
J. of Coastal Research	Erosion Hazard Vulnerability of US Coastal C	JCR_3	3	USA	2005	5,147
J. of Coastal Research	A Simple Method of Measuring Beach Profile	JCR_4	2	Portugal	2006	2,389
J. of Coastal Research	A New Global Coastal Database for Impact a	JCR_5	9	Greece; UK; Ireland; Germany; N	2008	4,788
J. of Coastal Research	Assessment of Vulnerability and Adaptation	JCR_6	1	Germany	2008	8,688
J. of Coastal Research	Sustainable Management of Surfing Breaks:	JCR_7	4	New Zealand	2009	11,534
J. of Coastal Research	Importance of Coastal Change Variables in E	JCR_8	3	USA	2010	4,311
J. of Coastal Research	The Healing Sea: A Sustainable Coastal Oce	JCR_9	2	Belgium	2009	11,155
J. of Coastal Research	Open-Ocean Barrier Islands: Global Influenc	JCR_10	2	USA	2011	7,538
J. of Coastal Research	Tracking Oil Slicks and Predicting their Traje	JCR_11	1	USA	2010	6,583
J. of Coastal Research	Classification of Coasts	JCR_12	1	USA	2004	8,476
J. of Coastal Research	Coastal Classification: Systematic Approach	JCR_13	1	USA	2004	20,515

Phase 1: Consultation on corpus design

Corpus linguists & academics from 12 departments of Civil Engineering consulted on design criteria:

- Peer-reviewed journals, preferably listed in Science Citation Index Expanded (SCI) or Social Sciences Citation Index (SSCI)
- Widely read & respected by researchers; considered “key journals” or “desired outlets for academic work”
- Higher impact factors (IF), 5-year IF, Eigenfactor, article influence (Thomson Reuters)
- Research articles selected by (a) Most cited; (b) Most viewed; (c) Most recent (1 article per volume)
- Minimum size of 1 million words recommended for specialised corpora (Kennedy 1998; Pearson 1998; Rea Rizzo 2010)

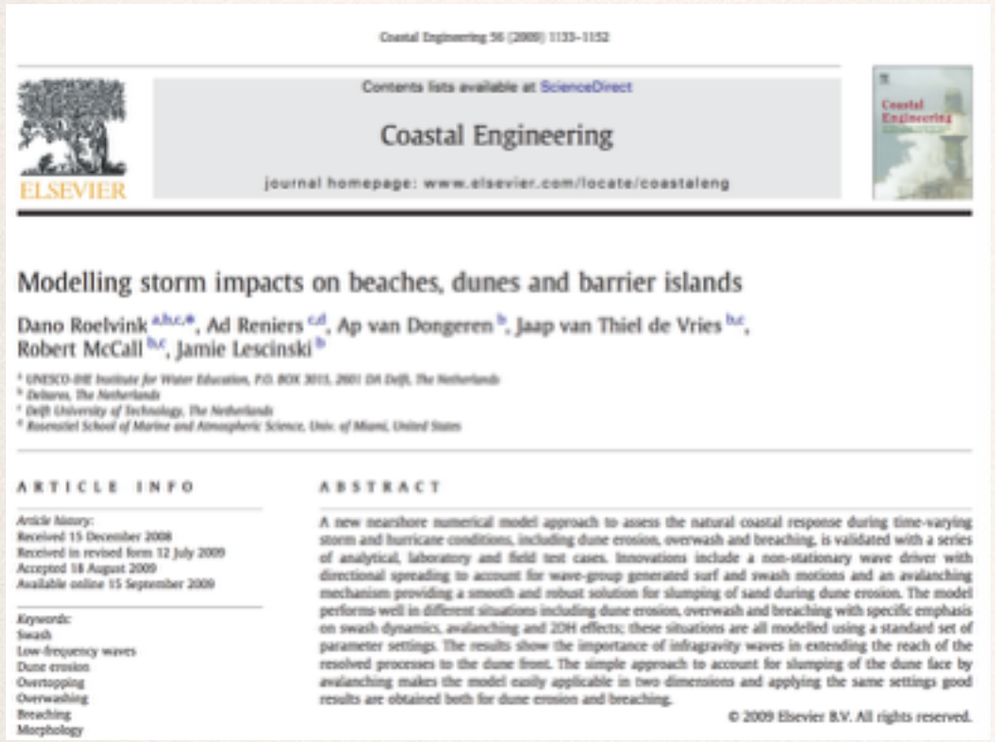
Phase 2: Construction of SCCERA

- HTML or PDF version of articles copied into MS Word
- Extraneous information removed (references, date of acceptance, author affiliation, contact info., tables & figures, equations)
- Text cleaned up using spelling & grammar checking function of MS Word (hyphenated words, conjoined words, character misreadings)
- HTML fragments (‘Table options’, ‘Turn Mathjax on’, etc.) removed using find & replace function in MS Word
- Articles saved as text-only (.txt) files

Phase 2: Construction of SCCERA

- 2nd round of cleaning up using text-only files (Greek symbols, etc.)
- Final document checked against original PDF file
- SCCERA part-of-speech (POS) tagged using CLAWS 4 (Lancaster University UCREL C7 tag set (Total no. tag types = 137): http://ucrel.lancs.ac.uk/claws7tags.html)

Phase 2: Construction of SCCERA



Coastal Engineering 56 (2009) 1133–1152

Contents lists available at ScienceDirect

Coastal Engineering

journal homepage: www.elsevier.com/locate/coastaleng

Modelling storm impacts on beaches, dunes and barrier islands

Dano Roelvink [a,b,c,*], Ad Reniers [c,d], Ap van Dongeren [b], Jaap van Thiel de Vries [b,c], Robert McCall [b,c], Jamie Lescinski [b]

[a] UNESCO-IHE Institute for Water Education, P.O. BOX 3015, 2601 DA Delft, The Netherlands
[b] Deltares, The Netherlands
[c] Delft University of Technology, The Netherlands
[d] Rosenstiel School of Marine and Atmospheric Science, Univ. of Miami, United States

ARTICLE INFO

Article history:
Received 15 December 2008
Received in revised form 12 July 2009
Accepted 18 August 2009
Available online 15 September 2009

Keywords:
Swash
Low-frequency waves
Dune erosion
Overtopping
Overwashing
Breaching
Morphology

ABSTRACT

A new nearshore numerical model approach to assess the natural coastal response during time-varying storm and hurricane conditions, including dune erosion, overwash and breaching, is validated with a series of analytical, laboratory and field test cases. Innovations include a non-stationary wave driver with directional spreading to account for wave-group generated surf and swash motions and an avalanching mechanism providing a smooth and robust solution for slumping of sand during dune erosion. The model performs well in different situations including dune erosion, overwash and breaching with specific emphasis on swash dynamics, avalanching and 2DH effects; these situations are all modelled using a standard set of parameter settings. The results show the importance of infragravity waves in extending the reach of the resolved processes to the dune front. The simple approach to account for slumping of the dune face by avalanching makes the model easily applicable in two dimensions and applying the same settings good results are obtained both for dune erosion and breaching.

UCREL CLAWS 7 Tagset

VVG = -ing participle of lexical verb

NN1 = singular common noun

NN2 = plural common noun

II = general preposition

```
^ <text>_NULL

^ Modelling_VVG storm_NN1 impacts_NN2 on_II beaches_NN2 ,_, dunes_NN2 and_CC
barrier_NN1 islands_NN2 Abstract_VV0@ A_ZZ1 new_JJ nearshore_NN1 numerical_JJ
model_NN1 approach_NN1 to_TO assess_VVI the_AT natural_JJ coastal_JJ
response_NN1 during_II time-varying_JJ storm_NN1 and_CC hurricane_NN1
conditions_NN2 ,_, including_II dune_NN1 erosion_NN1 ,_, overwash_NN1 and_CC
breaching_VVG ,_, is_VBZ validated_VVN@ with_IW a_AT1 series_NN of_IO
analytical_JJ ,_, laboratory_NN1 and_CC field_NN1 test_NN1 cases_NN2 ._.

^ Innovations_NN2 include_VV0 a_AT1 non-stationary_JJ wave_NN1 driver_NN1
with_IW directional_JJ spreading_NN1 to_TO account_VVI for_IF wave-group_JJ
generated_JJ@ surf_NN1 and_CC swash_VVI motions_NN2 and_CC an_AT1
avalanching_JJ@ mechanism_NN1 providing_VVG a_AT1 smooth_JJ and_CC robust_JJ
solution_NN1 for_IF slumping_VVG of_IO sand_NN1 during_II dune_NN1 erosion_NN1
._.

^ The_AT model_NN1 performs_VVZ well_RR in_II different_JJ situations_NN2
including_II dune_NN1 erosion_NN1 ,_, overwash_NN1 and_CC breaching_VVG
with_IW specific_JJ emphasis_NN1 on_II swash_NN1 dynamics_NN ,_,
avalanching_VVG and_CC 2DH_FO effects_NN2 ;_; these_DD2 situations_NN2 are_VBR
all_DB modelled_VVN using_VVG a_AT1 standard_JJ set_NN1 of_IO parameter_NN1
settings_NN2 ._.

^ The_AT results_NN2 show_VV0 the_AT importance_NN1 of_IO infragravity_NN1
waves_NN2 in_II extending_VVG the_AT reach_NN1 of_IO the_AT resolved_JJ@
processes_NN2 to_II the_AT dune_NN1 front_NN1 ._.

^ The_AT simple_JJ approach_NN1 to_TO account_VVI for_IF slumping_VVG of_IO
the_AT dune_NN1 face_NN1 by_II avalanching_VVG makes_VVZ the_AT model_NN1
easily_RR applicable_JJ in_II two_MC dimensions_NN2 and_CC applying_VVG the_AT
same_DA settings_NN2 good_JJ results_NN2 are_VBR obtained_VVN both_RR for_IF
dune_NN1 erosion_NN1 and_CC breaching_VVG ._.
```

Processing time (mins per RA)

Mean = 7.5

March 10
2014

April 19
2014

Processing time (mins per RA)

Common problems

1. Introduction

A primary goal of modeling physical processes in the atmospheric and hydrologic sciences is the prediction of a variable in time and/or space from a given set of inputs. How well a model fits the observed data (referred to as model evaluation, or sometimes as model validation) usually is determined by pairwise comparisons of model-simulated (or model-predicted) values with observations. Quantitative assessments of the degree to which the model simulations match the observations are used to provide an evaluation of the model's predictive abilities.

Frequently, evaluations of model performance utilize a number of statistics and techniques. Usually included in these tools are "goodness-of-fit" or relative error measures (bounded statistics, usually between 0.0 and 1.0) to assess the ability of a model to simulate reality. Often these statistics are based on the familiar Pearson's product-moment correlation coefficient (r) or its square, the coefficient of determination (R^2). These two statistics describe the degree of collinearity between the observed and model-simulated variates. They are almost always discussed in basic statistics texts and, consequently, are familiar to virtually all scientists. Unfortunately, both r and R^2 suffer from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of a model [*Willmott*, 1981; *Willmott et al.*, 1985; *Kessler and Neas*, 1994; *Legates and Davis*, 1997].

As knowledge of physical processes has increased, models have become more complex. Often these models include numerous parameters that are calibrated through optimization procedures, where a range in model parameters is sampled until the differences between the observed and model-simulated data are minimized [*Nash and Sutcliffe*, 1970; *Song and James*, 1991; *Hay*, 1998]. Stochastic calibration procedures are usually employed, which limits graphical analyses of scatterplots, for example, so that statistical analyses must be solely used. Consequently, statistics other than r and R^2 have been developed to describe better the degree of association between the observed and model-simulated data. The objectives of this paper are to (1) examine various goodness-of-fit measures and to identify limitations associated with each, and (2) suggest viable alternative measures for the evaluation of hydrologic and hydroclimatic models.

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

2. Statistics for Evaluation of Hydrologic and Hydroclimatic Models

In this paper, three basic methods for model evaluation will be discussed: the coefficient of determination R^2, the coefficient of efficiency E [*Nash and Sutcliffe*, 1970], and the index of agreement d [*Willmott et al.*, 1985]. In general, this paper addresses comparisons of model-simulated data (P) with the observed data (O) for the same set of conditions (i.e., a pairwise comparison) over a given time period divided into N time increments that can be of arbitrary duration (e.g., monthly or daily time steps).

2.1. Coefficient of Determination R^2

The coefficient of determination is the square of the Pearson's product-moment correlation coefficient (i.e., $R^2 = r^2$) and describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

1. Introduction
A primary goal of modeling physical processes in the atmo- spheric and hydrologic sciences is the prediction of a variable in time and/or space from a given set of inputs. How well a model fits the observed data (referred to as model evaluation, or sometimes as model validation) usually is determined by pairwise comparisons of model-simulated (or model-pre- dicted) values with observations. Quantitative assessments of the degree to which the model simulations match the obser- vations are used to provide an evaluation of the model's pre- dictive abilities.
Frequently, evaluations of model performance utilize a
number of statistics and techniques. Usually included in these
tools are "goodness-of-fit" or relative error measures (bound-
ed statistics, usually between 0.0 and 1.0) to assess the ability of
a model to simulate reality. Often these statistics are based on
the familiar Pearson's product-moment correlation coefficient
2
(r) or its square, the coefficient of determination (R). These
two statistics describe the degree of collinearity between the
observed and model-simulated variates. They are almost al-
ways discussed in basic statistics texts and, consequently, are
familiar to virtually all scientists. Unfortunately, both r and R
suffer from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of
a model [WiUmott, 1981; WiUmott et al, 1985; Kessler and Neas, 1994; Legates and Davis, 1997].
As knowledge of physical processes has increased, models have become more complex. Often these models include nu- merous parameters that are calibrated through optimization
Copyright 1999 by the American Geophysical Union.
Paper number 1998WR900018. 0043-1397/99/1998WR900018$09.00
procedures, where a range in model parameters is sampled
until the differences between the observed and model-
simulated data are minimized [Nash and Sutcliffe, 1970; Song
and James, 1991; Hay, 1998]. Stochastic calibration procedures
are usually employed, which limits graphical analyses of scat-
terplots, for example, so that statistical analyses must be solely
2
used. Consequently, statistics other than r and R have been
developed to describe better the degree of association between the observed and model-simulated data. The objectives of this paper are to (1) examine various goodness-of-fit measures and to identify limitations associated with each, and (2) suggest viable alternative measures for the evaluation of hydrologic and hydroclimatic models.
2. Statistics for Evaluation of Hydrologic and Hydroclimatic Models

Common problems

Words split with hyphens

1. Introduction

A primary goal of modeling physical processes in the atmospheric and hydrologic sciences is the prediction of a variable in time and/or space from a given set of inputs. How well a model fits the observed data (referred to as model evaluation, or sometimes as model validation) usually is determined by pairwise comparisons of model-simulated (or model-predicted) values with observations. Quantitative assessments of the degree to which the model simulations match the observations are used to provide an evaluation of the model's predictive abilities.

Frequently, evaluations of model performance utilize a number of statistics and techniques. Usually included in these tools are "goodness-of-fit" or relative error measures (bounded statistics, usually between 0.0 and 1.0) to assess the ability of a model to simulate reality. Often these statistics are based on the familiar Pearson's product-moment correlation coefficient (r) or its square, the coefficient of determination (R^2). These two statistics describe the degree of collinearity between the observed and model-simulated variates. They are almost always discussed in basic statistics texts and, consequently, are familiar to virtually all scientists. Unfortunately, both r and R^2 suffer from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of a model [*Willmott*, 1981; *Willmott et al.*, 1985; *Kessler and Neas*, 1994; *Legates and Davis*, 1997].

As knowledge of physical processes has increased, models have become more complex. Often these models include numerous parameters that are calibrated through optimization procedures, where a range in model parameters is sampled until the differences between the observed and model-simulated data are minimized [*Nash and Sutcliffe*, 1970; *Song and James*, 1991; *Hay*, 1998]. Stochastic calibration procedures are usually employed, which limits graphical analyses of scatterplots, for example, so that statistical analyses must be solely used. Consequently, statistics other than r and R^2 have been developed to describe better the degree of association between the observed and model-simulated data. The objectives of this paper are to (1) examine various goodness-of-fit measures and to identify limitations associated with each, and (2) suggest viable alternative measures for the evaluation of hydrologic and hydroclimatic models.

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

2. Statistics for Evaluation of Hydrologic and Hydroclimatic Models

In this paper, three basic methods for model evaluation will be discussed: the coefficient of determination R^2, the coefficient of efficiency E [*Nash and Sutcliffe*, 1970], and the index of agreement d [*Willmott et al.*, 1985]. In general, this paper addresses comparisons of model-simulated data (P) with the observed data (O) for the same set of conditions (i.e., a pairwise comparison) over a given time period divided into N time increments that can be of arbitrary duration (e.g., monthly or daily time steps).

2.1. Coefficient of Determination R^2

The coefficient of determination is the square of the Pearson's product-moment correlation coefficient (i.e., $R^2 = r^2$) and describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

Common problems

1. Introduction

A primary goal of modeling physical processes in the atmospheric and hydrologic sciences is the prediction of a variable in time and/or space from a given set of inputs. How well a model fits the observed data (referred to as model evaluation, or sometimes as model validation) usually is determined by pairwise comparisons of model-simulated (or model-predicted) values with observations. Quantitative assessments of the degree to which the model simulations match the observations are used to provide an evaluation of the model's predictive abilities.

Frequently, evaluations of model performance utilize a number of statistics and techniques. Usually included in these tools are "goodness-of-fit" or relative error measures (bounded statistics, usually between 0.0 and 1.0) to assess the ability of a model to simulate reality. Often these statistics are based on the familiar Pearson's product-moment correlation coefficient (r) or its square, the coefficient of determination (R^2). These two statistics describe the degree of collinearity between the observed and model-simulated variates. They are almost always discussed in basic statistics texts and, consequently, are familiar to virtually all scientists. Unfortunately, both r and R^2 suffer from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of a model [*Willmott*, 1981; *Willmott et al.*, 1985; *Kessler and Neas*, 1994; *Legates and Davis*, 1997].

As knowledge of physical processes has increased, models have become more complex. Often these models include numerous parameters that are calibrated through optimization procedures, where a range in model parameters is sampled until the differences between the observed and model-simulated data are minimized [*Nash and Sutcliffe*, 1970; *Song and James*, 1991; *Hay*, 1998]. Stochastic calibration procedures are usually employed, which limits graphical analyses of scatterplots, for example, so that statistical analyses must be solely used. Consequently, statistics other than r and R^2 have been developed to describe better the degree of association between the observed and model-simulated data. The objectives of this paper are to (1) examine various goodness-of-fit measures and to identify limitations associated with each, and (2) suggest viable alternative measures for the evaluation of hydrologic and hydroclimatic models.

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

2. Statistics for Evaluation of Hydrologic and Hydroclimatic Models

In this paper, three basic methods for model evaluation will be discussed: the coefficient of determination R^2, the coefficient of efficiency E [*Nash and Sutcliffe*, 1970], and the index of agreement d [*Willmott et al.*, 1985]. In general, this paper addresses comparisons of model-simulated data (P) with the observed data (O) for the same set of conditions (i.e., a pairwise comparison) over a given time period divided into N time increments that can be of arbitrary duration (e.g., monthly or daily time steps).

2.1. Coefficient of Determination R^2

The coefficient of determination is the square of the Pearson's product-moment correlation coefficient (i.e., $R^2 = r^2$) and describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

Words split with hyphens

Specialised words incorrectly identified as mistakes (collinearity)

Common problems

1. Introduction

A primary goal of modeling physical processes in the atmospheric and hydrologic sciences is the prediction of a variable in time and/or space from a given set of inputs. How well a model fits the observed data (referred to as model evaluation, or sometimes as model validation) usually is determined by pairwise comparisons of model-simulated (or model-predicted) values with observations. Quantitative assessments of the degree to which the model simulations match the observations are used to provide an evaluation of the model's predictive abilities.

Frequently, evaluations of model performance utilize a number of statistics and techniques. Usually included in these tools are "goodness-of-fit" or relative error measures (bounded statistics, usually between 0.0 and 1.0) to assess the ability of a model to simulate reality. Often these statistics are based on the familiar Pearson's product-moment correlation coefficient (r) or its square, the coefficient of determination (R^2). These two statistics describe the degree of collinearity between the observed and model-simulated variates. They are almost always discussed in basic statistics texts and, consequently, are familiar to virtually all scientists. Unfortunately, both r and R^2 suffer from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of a model [*Willmott*, 1981; *Willmott et al.*, 1985; *Kessler and Neas*, 1994; *Legates and Davis*, 1997].

As knowledge of physical processes has increased, models have become more complex. Often these models include numerous parameters that are calibrated through optimization procedures, where a range in model parameters is sampled until the differences between the observed and model-simulated data are minimized [*Nash and Sutcliffe*, 1970; *Song and James*, 1991; *Hay*, 1998]. Stochastic calibration procedures are usually employed, which limits graphical analyses of scatterplots, for example, so that statistical analyses must be solely used. Consequently, statistics other than r and R^2 have been developed to describe better the degree of association between the observed and model-simulated data. The objectives of this paper are to (1) examine various goodness-of-fit measures and to identify limitations associated with each, and (2) suggest viable alternative measures for the evaluation of hydrologic and hydroclimatic models.

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

2. Statistics for Evaluation of Hydrologic and Hydroclimatic Models

In this paper, three basic methods for model evaluation will be discussed: the coefficient of determination R^2, the coefficient of efficiency E [*Nash and Sutcliffe*, 1970], and the index of agreement d [*Willmott et al.*, 1985]. In general, this paper addresses comparisons of model-simulated data (P) with the observed data (O) for the same set of conditions (i.e., a pairwise comparison) over a given time period divided into N time increments that can be of arbitrary duration (e.g., monthly or daily time steps).

2.1. Coefficient of Determination R^2

The coefficient of determination is the square of the Pearson's product-moment correlation coefficient (i.e., $R^2 = r^2$) and describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

Common problems

1. Introduction

A primary goal of modeling physical processes in the atmospheric and hydrologic sciences is the prediction of a variable in time and/or space from a given set of inputs. How well a model fits the observed data (referred to as model evaluation, or sometimes as model validation) usually is determined by pairwise comparisons of model-simulated (or model-predicted) values with observations. Quantitative assessments of the degree to which the model simulations match the observations are used to provide an evaluation of the model's predictive abilities.

Frequently, evaluations of model performance utilize a number of statistics and techniques. Usually included in these tools are "goodness-of-fit" or relative error measures (bounded statistics, usually between 0.0 and 1.0) to assess the ability of a model to simulate reality. Often these statistics are based on the familiar Pearson's product-moment correlation coefficient (r) or its square, the coefficient of determination (R^2). These two statistics describe the degree of collinearity between the observed and model-simulated variates. They are almost always discussed in basic statistics texts and, consequently, are familiar to virtually all scientists. Unfortunately, both r and R^2 suffer from limitations that make them poor measures of model performance. Although these statistics continue to be used to determine how well a model simulates the observed data, they nevertheless provide a biased view of the efficacy of a model [*Willmott*, 1981; *Willmott et al.*, 1985; *Kessler and Neas*, 1994; *Legates and Davis*, 1997].

As knowledge of physical processes has increased, models have become more complex. Often these models include numerous parameters that are calibrated through optimization procedures, where a range in model parameters is sampled until the differences between the observed and model-simulated data are minimized [*Nash and Sutcliffe*, 1970; *Song and James*, 1991; *Hay*, 1998]. Stochastic calibration procedures are usually employed, which limits graphical analyses of scatterplots, for example, so that statistical analyses must be solely used. Consequently, statistics other than r and R^2 have been developed to describe better the degree of association between the observed and model-simulated data. The objectives of this paper are to (1) examine various goodness-of-fit measures and to identify limitations associated with each, and (2) suggest viable alternative measures for the evaluation of hydrologic and hydroclimatic models.

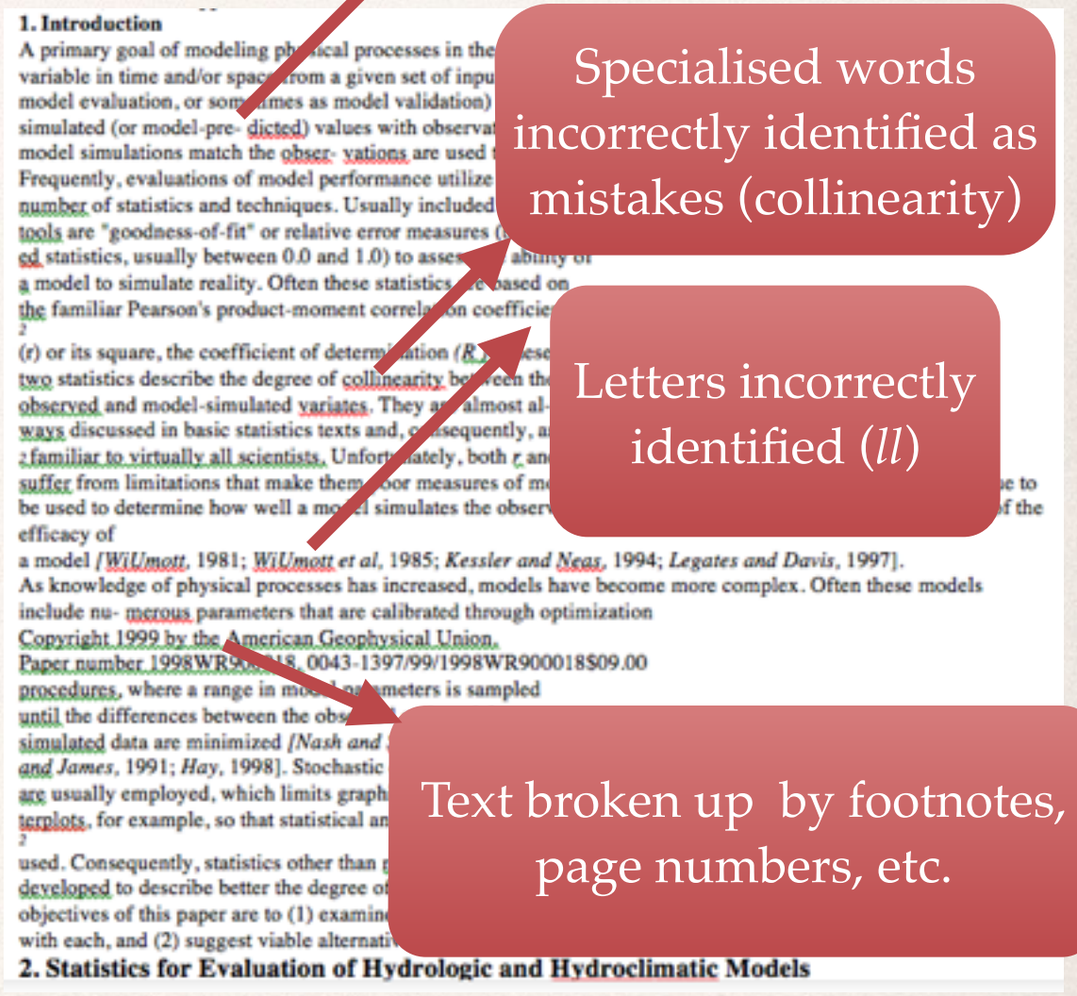
Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

2. Statistics for Evaluation of Hydrologic and Hydroclimatic Models

In this paper, three basic methods for model evaluation will be discussed: the coefficient of determination R^2, the coefficient of efficiency E [*Nash and Sutcliffe*, 1970], and the index of agreement d [*Willmott et al.*, 1985]. In general, this paper addresses comparisons of model-simulated data (P) with the observed data (O) for the same set of conditions (i.e., a pairwise comparison) over a given time period divided into N time increments that can be of arbitrary duration (e.g., monthly or daily time steps).

2.1. Coefficient of Determination R^2

The coefficient of determination is the square of the Pearson's product-moment correlation coefficient (i.e., $R^2 = r^2$) and describes the proportion of the total variance in the observed data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by



Common problems

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

served data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

$$R^2 = \left\{ \frac{\sum_{i=1}^{N} (O_i - \bar{O})(P_i - \bar{P})}{\left[\sum_{i=1}^{N} (O_i - \bar{O})^2\right]^{0.5} \left[\sum_{i=1}^{N} (P_i - \bar{P})^2\right]^{0.5}} \right\}^2 \quad (1)$$

where the overbar denotes the mean for the entire time period of the evaluation. Note, however, that the coefficient of determination is limited in that it standardizes for differences between the observed and predicted means and variances since it

adjusting factor would result in an increase in the correlation, possibly causing it to exceed 1.0 in extreme cases. Consequently, we do not advocate the use of such adjusting factors.

It should be noted that nonparametric or rank correlation methods also exist (e.g., Spearman's rho or Kendall's tau). As nonparametric statistics, they are less sensi… the data and generally provide a more robu… of the correlation between observed and pre… fortunately, rank correlation measures are… loss of information as interval/ratio data ar… dinal (ranked) form [see *Burt and Barber*, 19…

[…] and is given by (Equation 1) where x denotes…

and describes the proportion of the total variance in the ob- served data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

233

)

234

LEGATES AND MCCABE: EVALUATING "GOODNESS-OF-FIT" MEASURES

$_2 R=$

$$\mathbf{2\ (o,- o)(p,-p)}$$

$_{1=1|}$

$_N$

'

adjusting factor would result in an increase in the correlation, possibly causing it to exceed 1.0 in extreme cases. Conse- quently, we do not advocate the use of such adjusting factors.

It should be noted that nonparametric or rank correlation methods also exist (e.g., Spearman's rho or Kendall's tau). As nonparametric statistics, they are less sensitive to outliers in the data and generally provide a more robust characterization of the correlation between observed and predicted values. Un- fortunately, rank correlation measures are associated with a loss of information as interval/ratio data are converted to or- dinal (ranked) form [see *Burt and Barber*, 1996], and, like their parametric counterparts, they are not sensitive to additive and proportional differences between the observed and model- simulated values.

2.2. Coefficient of Efficiency *E*

The coefficient of efficiency *E* has been widely used to eval- uate the performance of hydrologic models [e.g., *Leavesley et al*, 1983; *Wdcox et al*, 1990]. *Nash and Sutcliffe* [1970] denned the coefficient of efficiency which ranges from minus infinity to 1.0, with higher values indicating better agreement, as

Common problems

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

served data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

$$R^2 = \left\{ \frac{\sum_{i=1}^{N} (O_i - \bar{O})(P_i - \bar{P})}{\left[\sum_{i=1}^{N} (O_i - \bar{O})^2\right]^{0.5} \left[\sum_{i=1}^{N} (P_i - \bar{P})^2\right]^{0.5}} \right\}^2 \quad (1)$$

where the overbar denotes the mean for the entire time period of the evaluation. Note, however, that the coefficient of determination is limited in that it standardizes for differences between the observed and predicted means and variances since it

adjusting factor would result in an increase in the correlation, possibly causing it to exceed 1.0 in extreme cases. Consequently, we do not advocate the use of such adjusting factors.

It should be noted that nonparametric or rank correlation methods also exist (e.g., Spearman's rho or Kendall's tau). As nonparametric statistics, they are less sensi… the data and generally provide a more robu… of the correlation between observed and pre… fortunately, rank correlation measures are… loss of information as interval/ratio data ar… dinal (ranked) form [see *Burt and Barber*, 19…

Mathematical symbols not recognised

[…] and is given by (Equation 1) where x denotes…

and describes the proportion of the total variance in the ob- served data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

233

)

234

LEGATES AND MCCABE: EVALUATING "GOODNESS-OF-FIT" MEASURES

2 R=

2 (o,- o)(p,-p) 1=1|

N

'

adjusting factor would result in an increase in the correlation, possibly causing it to exceed 1.0 in extreme cases. Conse- quently, we do not advocate the use of such adjusting factors.

It should be noted that nonparametric or rank correlation methods also exist (e.g., Spearman's rho or Kendall's tau). As nonparametric statistics, they are less sensitive to outliers in the data and generally provide a more robust characterization of the correlation between observed and predicted values. Un- fortunately, rank correlation measures are associated with a loss of information as interval/ratio data are converted to or- dinal (ranked) form [see *Burt and Barber*, 1996], and, like their parametric counterparts, they are not sensitive to additive and proportional differences between the observed and model- simulated values.

2.2. Coefficient of Efficiency *E*

The coefficient of efficiency *E* has been widely used to eval- uate the performance of hydrologic models [e.g., *Leavesley et al*, 1983; *Wdcox et al*, 1990]. *Nash and Sutcliffe* [1970] denned the coefficient of efficiency which ranges from minus infinity to 1.0, with higher values indicating better agreement, as

Common problems

Paper number 1998WR900018.
0043-1397/99/1998WR900018$09.00

served data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

$$R^2 = \left\{ \frac{\sum_{i=1}^{N} (O_i - \bar{O})(P_i - \bar{P})}{\left[\sum_{i=1}^{N} (O_i - \bar{O})^2\right]^{0.5} \left[\sum_{i=1}^{N} (P_i - \bar{P})^2\right]^{0.5}} \right\}^2 \quad (1)$$

where the overbar denotes the mean for the entire time period of the evaluation. Note, however, that the coefficient of determination is limited in that it standardizes for differences between the observed and predicted means and variances since it

adjusting factor would result in an increase in the correlation, possibly causing it to exceed 1.0 in extreme cases. Consequently, we do not advocate the use of such adjusting factors.

It should be noted that nonparametric or rank correlation methods also exist (e.g., Spearman's rho or Kendall's tau). As nonparametric statistics, they are less sensi- the data and generally provide a more robu of the correlation between observed and pre fortunately, rank correlation measures are loss of information as interval/ratio data ar dinal (ranked) form [see *Burt and Barber*, 19

[…] and is given by (Equation 1) where x denotes…

Mathematical symbols not recognised

Columns not recognised

and describes the proportion of the total variance in the ob- served data that can be explained by the model. It ranges from 0.0 to 1.0, with higher values indicating better agreement, and is given by

233

)

234

LEGATES AND MCCABE: EVALUATING "GOODNESS-OF-FIT" MEASURES

$_2 R=$

$$\mathbf{2\ (o,-\ o)(p,-p)}_{i=1]}$$

$_N$

adjusting factor would result in an increase in the correlation, possibly causing it to exceed 1.0 in extreme cases. Conse- quently, we do not advocate the use of such adjusting factors.

It should be noted that nonparametric or rank correlation methods also exist (e.g., Spearman's rho or Kendall's tau). As nonparametric statistics, they are less sensitive to outliers in the data and generally provide a more robust characterization of the correlation between observed and predicted values. Un- fortunately, rank correlation measures are associated with a loss of information as interval/ratio data are converted to or- dinal (ranked) form [see *Burt and Barber*, 1996], and, like their parametric counterparts, they are not sensitive to additive and proportional differences between the observed and model- simulated values.

2.2. Coefficient of Efficiency *E*

The coefficient of efficiency *E* has been widely used to eval- uate the performance of hydrologic models [e.g., *Leavesley et al*, 1983; *Wdcox et al*, 1990]. *Nash and Sutcliffe* [1970] denned the coefficient of efficiency which ranges from minus infinity to 1.0, with higher values indicating better agreement, as

Phase 3: Analysis of the corpus

Corpus linguistics literature

Student problems in academic writing

Research focus

Frequency & keywords analysis

Insights from civil engineers

Phase 3: Quantitative analysis of SCCERA

- Corpus analysis using WordSmith Tools 6.0 (Scott 2011)

- Comparisons across (a) RAs, (b) sub-sections, (c) sub-disciplines

- Word frequencies, keywords, key keywords, 2 to 8-word lexical bundles, type/token ratios, pedagogically significant concordance lines - e.g. disambiguation of near-synonymous words (Lee & Swales 2006)

Phase 3: Qualitative analysis of SCCERA

- Discourse analytical approach, investigating rhetorical characteristics of civil engineering RAs

- Move sequences in RA abstracts, introduction & discussion sections (often the most complex & problematic sections)

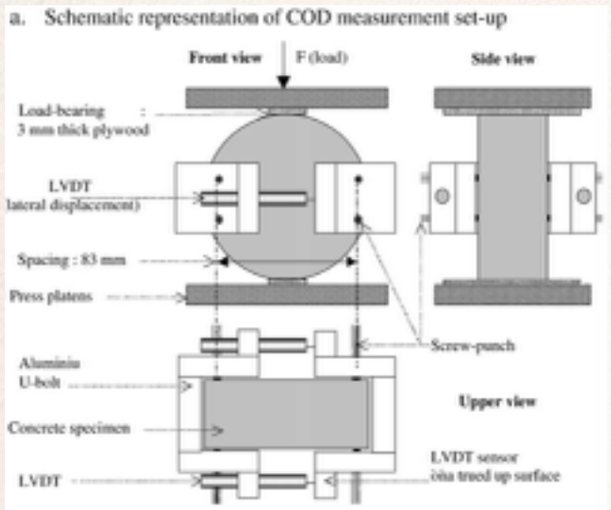
- Multimodality in civil engineering RAs

Word frequency (position)

et. al (24/5)	table (52)	models (89)	concrete (108)
model (29)	figure (59)	flow (93)	method (109)
fig (32)	may (61)	shown (92)	effects (112)
we (35)	values (64)	if (94)	mean (113)
between (36)	level (69)	case (95)	average (114)
time (37)	analysis (72)	large (97)	same (115)
used (39)	surface (76)	project (98)	stress (116)
results (44)	number (77)	area (100)	observed (117)
equation (45)	study (82)	effect (102)	change (126)
using (46)	value (83)	due (104)	see (192)

Multimodality in civil engineering RAs

(see Fig. 1a)



(see Fig. 1)

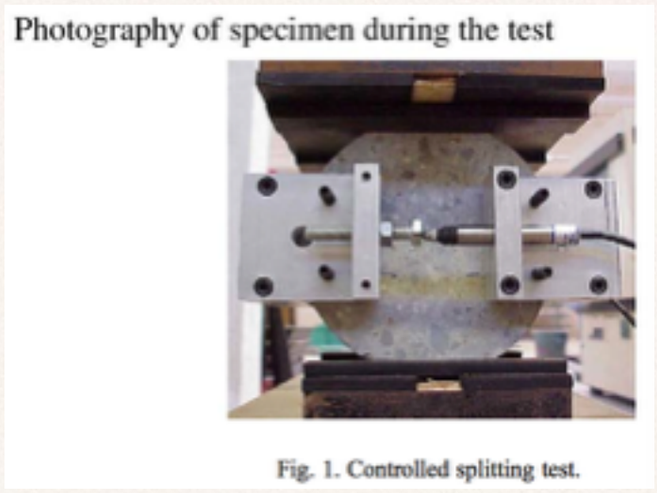


Fig. 1. Controlled splitting test.

[…], as in Eqn. (1):

$$J(x) = -D_e \frac{\partial c}{\partial x} + D_e \frac{zFE}{RTL} c + cv(x) \quad (1)$$

Fig. 3 presents…

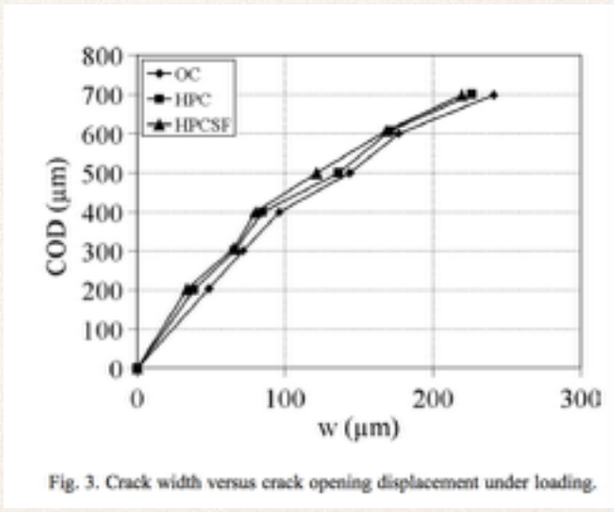


Fig. 3. Crack width versus crack opening displacement under loading.

(see Table 1)

Table 1
Details of test series and mix proportion

Mix ingredients (kg/m3)	OC	HPC	HPCSF
Coarse aggregate, 12.5–20 mm	777	550	579
Medium agrregate, 4–12.5 mm	415	475	465
Sand (Boulonnais), 0–5 mm	372	407	442
Sand (Seine), 0–4 mm	372	401	435
Cement CPA-CEM I 52.5	353	461	360
Silica fume	–	0	22
SP (e.s.)	–	12.4	12
Retarder (e.s.)	–	3.3	2.5
Total water	172	146	136
w/c	0.49	0.32	0.38
w/(*C*+SF)	0.49	0.32	0.36

(Fig. 2)

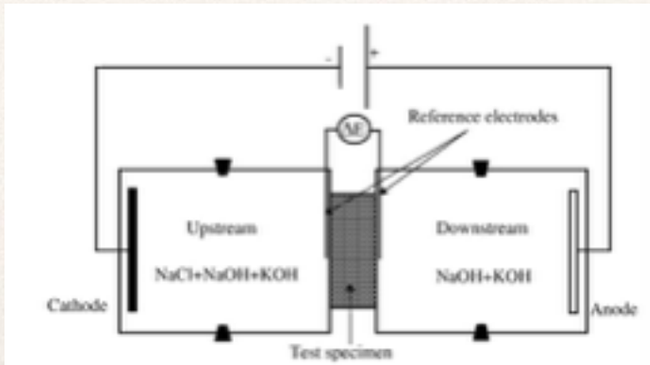


Fig. 2. Migration cell.

Epistemic language

Modal Verbs Could Couldn't May Might Should Shouldn't Would Wouldn't Will Won't **Adjectives** Always Apparent Certain A certain extent Clear Evident Possible Probable **Nouns** Claim Doubt Estimate Evidence Possibility **Lexical Verbs** Appear Argue Assume Believe Claim Doubt Estimate Expect Indicate Know Predict Presume Propose Seem Speculate Suggest Suppose Tend Think **Adverbs** About Actually Almost Apparently Approximately Around Certainly Clearly Definitely Doubtless Essentially Evidently Frequently Generally In fact Indeed Largely Likely Never Normally Obviously Of course Often Perhaps Possibly Presumably Probably Quite Rarely Relatively Sometimes Surely Undoubtedly Usually

Epistemic language

Modal Verbs Could Couldn't May Might Should Shouldn't Would Wouldn't Will Won't **Adjectives** Always Apparent Certain A certain extent Clear Evident Possible Probable **Nouns** Claim Doubt Estimate Evidence Possibility **Lexical Verbs** Appear Argue Assume Believe Claim Doubt Estimate Expect Indicate Know Predict Presume Propose Seem Speculate Suggest Suppose Tend Think **Adverbs** About Actually Almost Apparently Approximately Around Certainly Clearly Definitely Doubtless Essentially Evidently Frequently Generally Indeed Largely Likely Never Normally Obviously Of course Often Perhaps Possibly Presumably Probably Quite Rarely Relatively Sometimes Surely Undoubtedly Usually

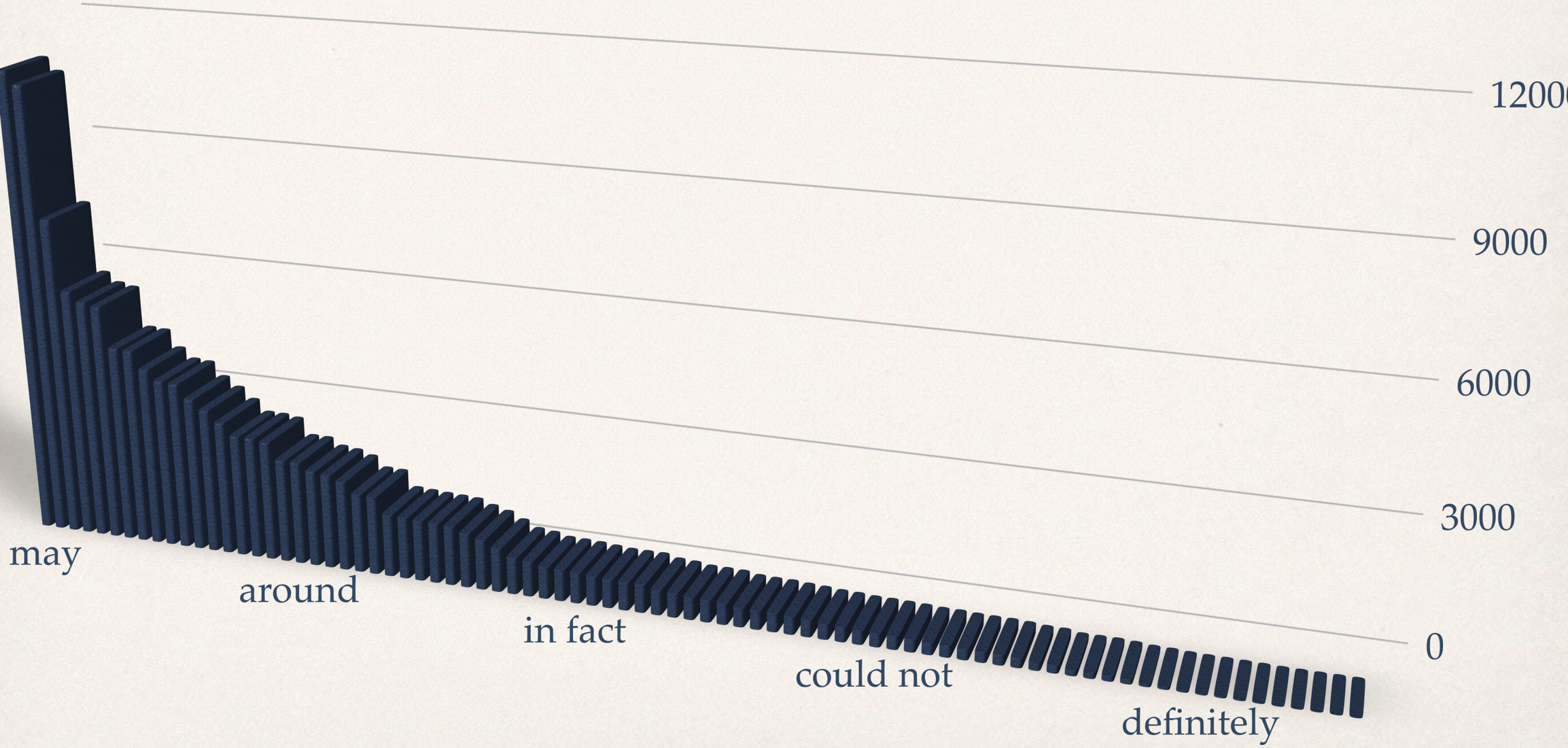
"Modes of knowing": Communicating doubts, certainties & guesses

Epistemic language

“Modes of knowing”: Communicating doubts, certainties & guesses

Most frequent epistemic items in academic writing (Hyland & Milton 1997)

Most frequent epistemic items in academic writing (Hyland & Milton 1997)



Epistemic items in SCCERA (frequency)

may (11,127)	could (4,044)	appear (2,267)	certain (1,204)
estimate(s) (10,823)	possible (3,713)	approximately (2,160)	quite (1,041)
will (7,703)	expect (3,488)	evidence (1,902)	argue (875)
about (6,019)	predict (3,223)	might (1,889)	indeed (833)
indicate(s) (5,794)	estimate (N) (2,942)	tend (1,502)	apparent (757)
would (5,722)	likely (2,934)	clear (1,478)	wouldn't (9)
should (4,754)	relatively (2,884)	seem (1,456)	won't (6)
assume (4,727)	often (2,486)	usually (1,439)	couldn't (3)
suggest (4,315)	around (2,480)	almost (1,419)	doubtless (3)
propose (4,074)	generally (2,311)	clearly (1,302)	shouldn't (2)

Modal expressions

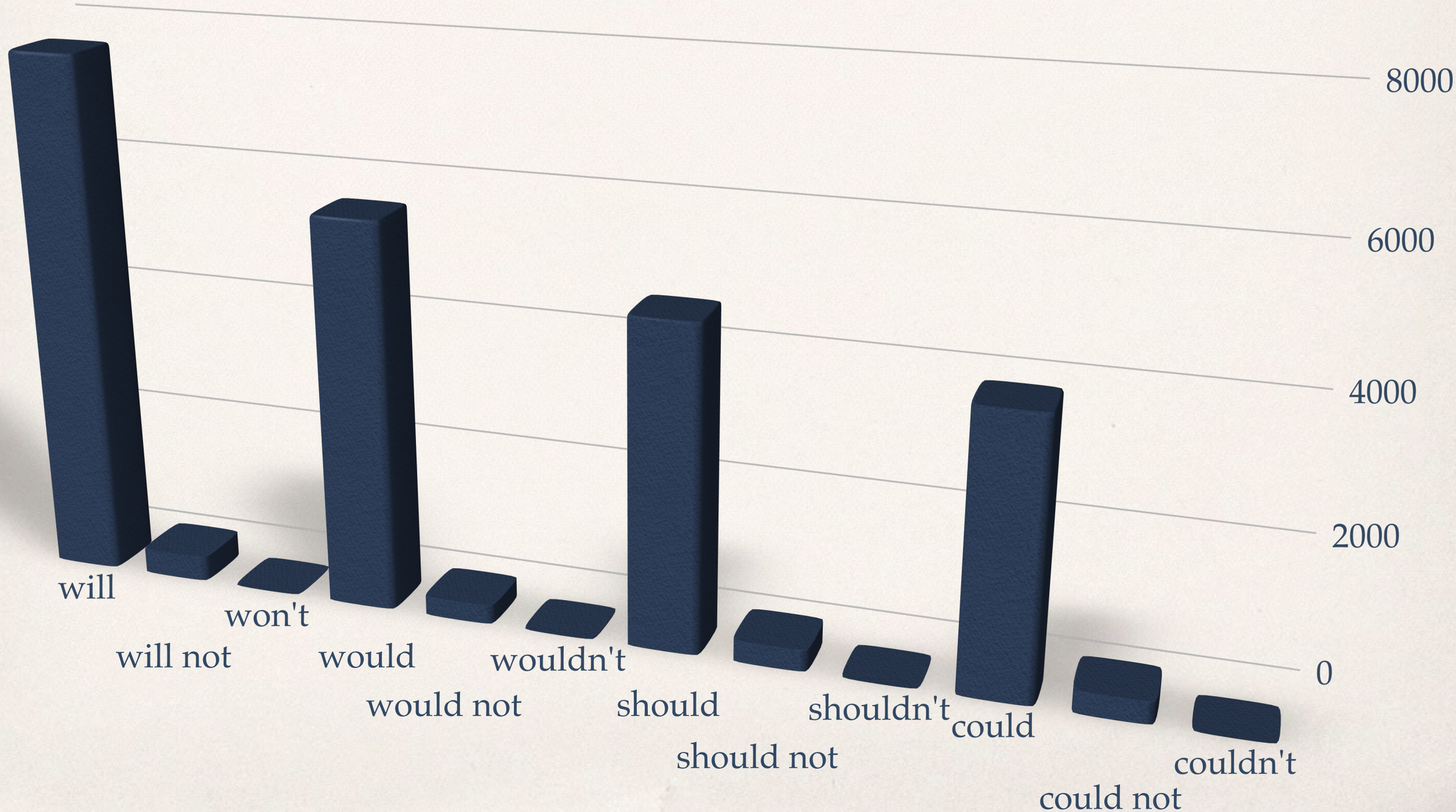
will not would

would not should

should not

could not

Modal expressions



POS - CLAWS tagset

Position	POS tag	Info.	SCCERA	Medical	Brown
1	N	nouns	32.2%	29.1%	23.1%
2	V	verbs	13.4%	11.1%	15.5%
3	I	prepositions	13.4%	—	—
4	J	adjectives	10.2%	9.7%	6.9%

Coastal Engineering: Keywords vs. SCCERA (position)

wave (1)	storm (13)	numerical (23)	reef (33)
sea (2)	shoreline (14)	height (24)	waters (34)
coastal (3)	coast (15)	bed (25)	salinity (35)
ice (4)	erosion (16)	islands (26)	breakwater (37)
waves (5)	tidal (17)	water (27)	surge (38)
ocean (6)	tide (18)	shore (28)	swash (39)
breaking (7)	beaches (19)	offshore (29)	Atlantic (40)
beach (9)	currents (20)	island (30)	coasts (41)
shelf (10)	depth (21)	dune (31)	figure (42)
wind (12)	arctic (22)	runup (32)	shelves (44)

Coastal Engineering: Keywords vs. BNC (position)

wave (2)	ocean (14)	breaking (25)	tsunami (36)
et al (3/4)	water (16)	wind (26)	eq (37)
coastal (5)	figure (17)	tidal (27)	boundary (38)
ice (6)	velocity (18)	flow (28)	erosion (39)
model (8)	surface (19)	beach (29)	the (40)
fig (9)	depth (20)	height (30)	values (41)
equation (10)	numerical (21)	storm (32)	elevation (42)
sea (11)	sediment (22)	level (33)	measurements (43)
waves (12)	shelf (23)	measured (34)	salinity (45)
data (13)	shoreline (24)	results (35)	simulation (46)

Keywords: Hard vs. soft sub-disciplines of civil engineering

Hard (Mechanics & Structures)		**Soft** (Infra- structure dvlpt)	
damping	load	project	risk
response	steel	construction	pavement
beam	force	management	risks
structural	displacement	projects	team
stiffness	strain	cost	research
bridge	equation	success	safety
control	damage	life	leadership
vibration	frequency	costs	performance
damper	excitation	managers	process